Modality in New Englishes:

A Corpus-Based Study of Obligation and Necessity

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

vorgelegt von

Gabriela Veronica Diaconu

aus Braşov

SS 2011

Erstgutachter: Prof. Dr. Dr. h.c. Christian Mair

Zweitgutachter: Prof. Dr. Bernd Kortmann

Vorsitzender des Promotionsausschusses
der Gemeinsamen Kommission der
Philologischen, Philosophischen und Wirtschafts-
und Verhaltenswissenschaftlichen Fakultät: Prof. Dr. Hans-Helmuth Gander
Datum der Fachprüfung im Promotionsfach: 24.01.2012

Contents

List of tables

List of figures

List of abbreviations

AmE	American English
AusE	Australian English
BNC	British National Corpus
BrE	British English
Brown	Brown Corpus of American English
BSAfE	Black South African English
COLT	The Bergen Corpus of London Teenage Language
CookE	Cook Island English
EFL	English as a foreign language
EModE	Early Modern English
ENL	English as a native language
ESD	English as a second dialect
ESL	English as a second language
FijE	Fiji English
FLOB	The Freiburg – LOB Corpus of British English
FRED	Freiburg Corpus of English Dialects
Frown	The Freiburg – Brown Corpus of American English
HKE	Hong Kong English
KenE	Kenya English
ICE	International Corpus of English
ICE-GB	International Corpus of English – Great Britain
ICE-HK	International Cropus of English – Hong Kong
ICE-India	International Corpus of English – India
ICE-Ireland	International Corpus of English – Ireland
ICE-JA	International Corpus of English – Jamaica
ICE-SAfE	International Corpus of English – South Africa
ICE-T&T	International Corpus of English – Trinidad & Tobago
IndE	Indian English
IrE	Irish English
JamE	Jamaican English
JC	Jamaican Creole

LLC	The London-Lund Corpus of Spoken English
LModE	Late Modern English
LOB	The Lancaster-Oslo/Bergen Corpus of British English
ME	Middle English
ModE	Modern English
NZE	New Zealand English
OE	Old English
PhilE	Philippine English
SamE	Samoa English
SEU	The Survey of English Usage
SingE	Singapore English
SBC	Santa Barbara Corpus of Spoken American English
SOV	Subject-Object-Verb
SVO	Subject-Verb-Object
TorE	Toronto English
WSAfE	White South African English

Acknowledgements

This dissertation would not have been possible without the help of so many people. First and foremost, I wish to thank and express my deepest appreciation to my supervisor Christian Mair for his constant support and patience in the process of finishing this thesis. Without his help, expertise in corpus linguistics and the study of change in present-day English, guidance and enthusiasm this thesis would not have been possible. I would like to thank my co-advisor Bernd Kortmann for his constructive criticism and input on individual aspects of the thesis. I am indebted to both of them for introducing me to Linguistics and to the principles of research in this field of study.

I gratefully acknowledge the financial support for this project which was provided by the International Graduate Academy at the Albert-Ludwig University in Freiburg by way of a scholarship of the State Law on Graduate Funding Baden-Würrtemberg (LGFG) from May 2008 – October 2010. I would like to thank Sylvie Nantcha for her help and indispensable advice in the application procedure. I would like to thank the Headmaster of the Albert-Ludwig University, Hans-Jochen Schiewer, for providing me a financial aid in the final stages of my dissertation project. I am grateful for a travel grant offered by the Hermann Paul School for Languages which enabled me to participate at the conference "CHRONOS 10: 10th International Conference on Tense, Aspect, Modality and Evidentiality" at Aston University Birmingham, UK (April 2011).

I thank Benedikt Szmrecsanyi for introducing me to R and for his expert knowledge on the multivariate logistic regression with corpus linguistic data. For proofreading several drafts and for their suggestions and detailed comments on my writing and on various aspects of my research I thank Nicholas Smith and Rachel Hill.

I am indebted to Nicholas Smith and Geoffrey Leech for giving me the opportunity to undertake research work placement at the University of Salford, England, in January 2010 as related with the project "Language change in news and science at the turn of the 20th century: a corpus-based study".

For their help, advice and encouragement I owe a lot to my former colleagues at the English Department at the University of Freiburg, and especially to: Mădălina Chitez, Susanne Gundermann, Glenda Leung, Andrea Moll and Friederike Seggewiß. I owe many thanks to Friederike Seggewiß for her willingness to share with me her knowledge and ideas about modality and modal auxiliaries in present-day English as well as for her assistance in many aspects of my data analysis.

In addition, thanks go to the audiences of the several conferences where I was given the opportunity to present my research. Needless to say, in spite of all the help received in writing this thesis all inconsistencies and shortcomings are entirely my own.

The process of writing a thesis can be both arduous and enjoyable, and many friends have contributed to my well-being during this time. I would like to thank my good friends Mădălina Chitez and Octavian Iercan for their help and advice which has reached

beyond matters concerning only academic life in Freiburg, and for making me feel less homesick at times. My heartfelt thanks go to Dorina Descas and Katarzyna Matraszek for their sincere friendship and for sharing with me their experience and wisdom. This list would not be complete without mentioning my friends Alia Câlțea, Ionela Deca, Antonia Efron, Bianca Harnisch, Viorica Popa, who, in spite of the distance which has separated us, have always encouraged me. Thank you for being such good friends!

The most valuable support has been offered by my parents. This thesis is dedicated to my parents, who have given me their unconditional support throughout my life and believed in all my efforts.

Last but not least, I wish to thank my husband, Miguel, who has encouraged me and stood by me all this time. He is the source of my happiness and I feel very fortunate to have him in my life.

1 Introduction

1.1 Modal usage and change in progress in new standard Englishes: an overview

Ongoing changes in the English modality system have been described mainly on the basis of British and American written texts in previous real-time corpus-based studies (Krug 2000; Leech 2003; Smith 2003). According to these studies, the modality system of present-day English grammar is an area which has undergone significant shifts in the frequency, distribution and semantic developments of modals, and the so-called semi-modals. The contrast in meaning and function between root/non-epistemic as well as epistemic obligation and necessity seems to be especially salient.

General tendencies show that the root meaning of the modal *must* is decreasing quantitatively, while semi-modals are rapidly increasing in frequency. Similarly, epistemic *must* has increased in frequency especially in British English, whereas, *have to* seems to have maintained its prototypical use as a marker of objective necessity. More recent studies (Leech et al. 2009: 109) show that the semi-modal *have to* has undergone a further development towards epistemic use, which is still a rare phenomenon in British and American written texts. By contrast, the semi-modal *(have) got to* is usually associated with spoken English in informal contexts while semi-modal *need to*, a recent development in the modality system, expresses unmarked necessity (Leech et al. 2009: 111).

As the amount of research in describing postcolonial varieties of English is growing constantly, it seems necessary to complement these studies with analyses of spoken material which document the extent of related innovations of language-internal variation patterns in Standard English world-wide. To this end, the present thesis presents an in-depth analysis of recent changes occurring with the four commonly used expressions of strong obligation and necessity, i.e. *must*, *have to*, *(have) got to* and *need to* in three newly standardized or standardizing ex-colonial varieties of English: Jamaican English (JamE), Indian English (IndE), and Irish English (IrE). Since spoken language, in contrast with the stable written norm, is considered “the driving force of change” (Mair & Leech 2006: 328) in grammar, and in particular in modality, the present study aims to find out if these expressions follow distinct paths of development

in the previously mentioned varieties as compared to British and American English. Additionally, emphasis is placed on the relevance of style (formal vs. informal) in ongoing changes of modality.

Since instances of language change are initiated in language use, the study adopts a functionalist usage-based view on the modal – semi-modal alternation as a form-meaning relationship (e.g. Croft 2000; Krug 2000). In particular, the empirical observations concerning inter-regional differences of modal usages are viewed not only in terms of the processes of language contact and diffusion but also of convergence. The present study focuses on patterns of convergence and divergence in the use of modal obligation/necessity in spoken interactions in present-day English. It aims to uncover insights into the overall dynamics of the notion of standard norm (British or American) through the exploration of the discourse frequency and semantic distinctions of *must*, *have to*, *(have) got to* and *need to*. The main research approach consists of integrating findings from analyses undertaken in real-time and apparent-time paradigms to compare diachronic and synchronic variation patterns in modality (see section 1.2).

Mapping language-internal variation in JamE, IndE and IrE would not have been possible without a corpus investigation. Building on existing work on written language (e.g. on data from BNC, Brown, Frown, LOB, and FLOB), the descriptive analysis of regionally and stylistically diverse spoken material retrieved from the recently completed Jamaican, Indian and Irish components of the *International Corpus of English* (ICE) will give insight into ongoing change in the apparent-time paradigm in the New Englishes. A systematic analysis of the co-occurrence of variation patterns in the New Englishes has to consider several dimensions: regional, social, stylistic, and medium (e.g. printed vs. non-printed writing, writing vs. speech). In this respect, the parallel sampled sub-corpora of ICE, representing a large number of English-speaking world regions, are the most appropriate resources for such a study.

Despite the underlying uniformity of these varieties, more recent research suggests a reduction in the well-documented British influence in the outer circle varieties, which has favoured global exonormative influence, for example from colloquial American speech-forms (on the emergence of written Standard Jamaican English, see Mair 2002a). Conversely, the various 'local' influences found in individual varieties contribute to their 'indigenisation' which may proceed faster in speech than in

writing. Typical for these communities is the coexistence of international (or standard) and 'local' (or non-standard) usages, with variation functioning as a stylistic resource. Although they mostly share a British heritage, New Englishes often display tension between the vernacular and the colonial standard (Bhatt 2004: 1016). This tension raises issues about the legacy of colonial English (Hickey 2004d), as each variety has developed "distinct linguistic ecologies – their own contexts of function and usage" (Kachru 1986: 1). This will be shown to be true when looking at locally biased distributions and functions of the selected modality markers.

The general assumption is that, as in the case of modals, the emergence of semi-modals is very common in varieties of English all over the world (cf. 'common core or nucleus' in Nelson 2006: 129; Quirk et al. 1985: 16), but with slight variations in their realizations. From this perspective, variation is not just an instance of formal changes on the morphosyntactic level, but is motivated by the interaction between language-internal factors and the pragmatic context of discourse, which enables functional alternation between modal items. To compare variation patterns of the modality system in samples from three different varieties both quantitative and qualitative findings are integrated in the research design.

Last but not least, the present study seeks to provide further evidence relevant to the recurring debate about 'angloversals', i.e. the claim that the New Englishes display common developmental trends, which can be observed regardless of their particular history of contact and diffusion (Mair 2003; Sand 2005).

The present study is organised into eight chapters. Chapters 2 and 3 lay down the theoretical foundations in regards to linguistic research on the New Englishes. A brief introduction of the relevant terminology within the study of the so-called New Englishes will help underline the relationship between native, non-native and second-language varieties. The chapter ends with a synthesis of the main trends in linguistic research on the New Englishes. Chapter 3 focuses on the historical and linguistic stages in the emergence of Standard English in Ireland, Jamaica and India. The purpose of this chapter is to emphasise both the common British heritage as well as the different fates in the development of these varieties, significantly that they pursue the same goal: the recognition as emerging standard national varieties.

The theoretical and methodological framework is introduced in Chapter 4. A detailed review of the most relevant corpus-based investigations will integrate our topic into the current studies of modal variation and change in present-day English. In addition, the chapter provides a detailed description of the database under analysis. The remaining part of the chapter consists of establishing a set of relevant methodological steps with regard to the data retrieval and analysis.

The main body of the thesis consists in a systematic quantitative and qualitative corpus-based analysis of the alternations between modal *must* and the semi-modals *have to*, *(have) got to* and *need to*, which is the subject of Chapters 5 to 7. To begin with, Chapter 5 focuses on the overall distribution of forms bearing all the morphosyntactic realizations of the four items across the selected text categories in ICE. Discourse frequency is considered to be a relevant determinant of variation and change. To explain synchronic regional and stylistic distribution patterns, the quantitative analysis draws on diachronic evidence from the relevant literature (e.g. Krug 2000).

In Chapter 6, the quantitative evidence on the use of the four markers is substantiated with a detailed qualitative examination of their root meanings. The aim of such an examination is to identify recurrent paths of change within root necessity in the light of the objective vs. subjective dichotomy, which contribute to the *subjectification* of modal markers in present-day English.

Chapter 7 extends the descriptive approach to include variationist techniques, which are common in the sociolinguistics study of variation and change. The purpose of such a "descriptive-interpretative" approach is to quantify several language-internal and language-external variables and to identify shifts within the constraints of obligation/ necessity, which would allow us to assume ongoing change in the system.

Finally, Chapter 8 provides a summary of the major findings in the light of the current research on New Englishes, and indicates some potential directions of future research.

1.2 A corpus-based study of modal variation and change in progress

The structural and semantic properties shared by modals and semi-modals in present-day English are rooted in the grammaticalization process in which these markers have undergone a transformation from pre-modal to modal features of necessity (Hopper & Traugott 2003; Krug 2000). Modal behaviour is often referred to as "puzzlingly complex" (Warner 1993: xi), and studies which highlight the asymmetric discourse frequency and semantic distribution of modal *must* and semi-modal[1] *have to* in different English speaking communities describe them as both 'recessive' and 'emergent' (cf. Jacobsson 1994), or even 'on the move' (Leech 2003).[2] From regional perspective, it is in AmE that semi-modals are currently spreading most rapidly at the expense of modals (Leech 2003: 237; Mair & Leech 2006: 328).

An increasing number of recent corpus-based investigations are available for newer and newly standardized varieties of English such as, e.g. AusE and NZE (Collins 2005; 2007; 2008; 2009a; 2009b), IndE (Leitner 1991; Schmied 1994; Shastri 1988; Wilson 2005), CanE (Dollinger 2008; Tagliamonte & D'Arcy 2007), South Pacific Englishes (Biewer 2009), and, most recently, SingE (Bao 2010). All these studies set off to identify trends of register-specific variation patterns in the distribution of modals and semi-modals in ex-colonial Englishes while assessing their development in light of their 'colonial lag', compared to BrE and AmE. The findings indicate increased exonormative linguistic influence in the form of American colloquial speech habits, e.g. the use of semi-modal *have to*. At the same time, local innovations in speech in the form of regional and stylistic variation patterns – often the result of contact-induced change through grammaticalization – are increasingly recognized.[3]

While many of these studies mainly report frequency patterns of modal items, they show less concern for an integrative approach of the ongoing semantic change.

[1] Despite the abundance of terminology for this verbal category (see also section 4.6.2), the term 'semi-modal' will be used throughout the whole study. Any diverging terminology will be used as linked to a specific author, unless otherwise indicated.

[2] See also Quirk et al. (1985: 220) on the different uses of modals found between native varieties of English.

[3] More recently, van Rooy & Rossouw (2011) and Rossouw & van Rooy (2011) have analysed the development of modality markers in South African English in data from the still incomplete ICE-SAfE component as compared to British English and stress the influence coming from Afrikaans and Xhosa-English.

Generally, correlations between semantic, morphosyntactic and pragmatic features as associated with modal necessity are studied mainly in present-day standard and non-standard BrE and AmE: e.g. Coates (1983), Depraetere & Verhulst (2008), Tagliamonte (2004), Tagliamonte & Smith (2006) and Tagliamonte & D'Arcy (2007). The differences between these features usually concern the strong core of the "obligation/necessity cluster" (Coates 1983: 31) and are treated in terms of the opposition root/non-epistemic vs. epistemic modality or subjective vs. objective modality, respectively (e.g. Coates 1983; Jacobsson 1979; Palmer 1990 [1979]; Perkins 1983; Tregidgo 1982; Westney 1995).

Currently, in standard educated JamE and IrE the use of obligation/necessity is little studied.[4] Among the notable exceptions are Mair's (2009b) quantitative corpus-based examination of obligation/necessity in ICE-JA and Corrigan's (2000) study on the non-standard vernacular use of such modal expressions in an Northern Irish English dialect.

Against this background, the present study examines possible changes within the strong modality system in three newer varieties of standard Englishes. To this end, spoken standard JamE, standard IndE and standard IrE will be analysed to gain insights on the development of obligation/necessity in discourse.

To introduce the subject-matter, the examples from (1) to (10) selected from the spoken part, i.e. the private and public dialogues, of the Jamaican, Indian and Irish components of ICE express 'strong' obligation/necessity (Bybee et al. 1994; Smith 2003). Each variant – *must*, *have to* /*'ve to*, *(have) got to*/*'ve got to* or *need to* – communicates diverse semantic and stylistic realizations of either internally (subjective) or externally (objective) motivated strong obligation/necessity:

(1) <#>When I read your thesis statement I **must** literally have an idea of what this essay is about <#>The thesis statement **must** give direction and focus to your essay<,> (ICE-JA.S1B-005.txt)

[4] More recently, Deuber (2010) studied the modality markers *can*/*could* and *will*/*would* in a related Caribbean variety from the Republic of Trinidad and Tobago by examining the still incomplete ICE-T&T component. Likewise, Ronan (2011) analyses the modals *would* and *will* on the basis of data from ICE-Ireland. Further non-standard Northern Irish English use of modal markers is documented in Tagliamonte (2004) and Tagliamonte & Smith (2006).

(2) Mm it seems strange <#> I **have to** take Julie down at her usual time and Chloe starts at ten past nine <,> so I 'll just you know there 's no point in me doing anything else going anywhere else (ICE-Ireland.S1A-001.txt)

(3) <S1A-070$B> <#> I**'ve to** stay in and do this project and then I 'm going to watch Peter 's Friends <S1A-070$A> <#> Yeah me too <#> Take off my boots and dry my feet <S1A-070$ (ICE-Ireland.S1A-070.txt)

(4) <$A><#>But Sir that **has got to** be the court of first resort or we are going to end up with a society where children see violence on the part of teachers and violence on the part of parents and they believe that that is a way just to go (ICE-JA.S1A-096.txt)

(5) She**'s** </w> **got to** work from <,,> nine <w> O'clock </w> till five <w> O'clock </w> <ICE-IND:S1A-021#128:1:C> And there are no tables <,> no chairs and she **has to** stand (ICE-India.S1A-021.txt)

(6) In Madras to go and see a film <,,> we **need to** spend a lot of money <,> for per persons <,> <{> <[> yes <,> ten <indig> rupees </indig> fifteen <indig> rupees </indig> <$C>(ICE-India.S1A-024.txt)

(7) Because some <,> students come into the schools <{1> <[1> knowing a language </[1> which is not English <#> </{2> <[2> And </[2> they **need to** understand it as a second language <{3> <[3> so as </[3> to be able to learn it in school (ICE-JA.S1A-001.txt)

(8) <#>Our organization looks at education preventive action<,,> aimed at that because <}><->we we</-> <=>we</=></}> believe that there **has to** be a solution <#>Persons cannot just take off leave children to basically follow the same path<,> (ICE-JA.S1B-040.txt)

(9) And there 's a <,> Derry 's wine and cheese party 's on next Saturday <{1> <[1> <#> So </[1> Jane has tickets to sell but I says I don't know how I got out of that for I got sent raffle tickets but they didn't send me <{2> <[2> tickets for the <,> so I **have to** buy </[2> <,> because they **'ve to** be in for that <{3> <[3> Saturday </[3> so we 're all going to go to that next <{4> <[4> Saturday </[4> (ICE-Ireland. S1A-012.txt)

(10) <I><$A><#>Mr <@>Pepper</@> you **have to** speak very loudly <#>The judge and everybody **must** hear you (ICE-JA.S1B-068.txt)

In the examples (1), (2), and (3) necessity is reported through the speaker, however the origin of the source differs in each case. In the first instance in (1) the source of necessity with modal *must* stems from the speaker, who expresses his beliefs as "self-obligation" (Leech 1987: 77) or as "insistent self-exhortation" (Collins 2005: 252), such as the necessity to capture the main idea of a thesis statement in a few lines of written

text. For this reason the modal in (1) is a good instance of prototypical root modality. In the second instance from the same example, the necessity expressed with *must* appears to be external to the speaker by referring to a prerequisite in academic writing. In examples (2) and (3) the necessity appears to be conveyed by the speaker's own will as self-exhortation. However, a careful reading shows that the necessity with semi-modal *have to* and its contracted form *'ve to* is motivated by external circumstances to the speaker, such as required by a strict schedule in (2) or a possible deadline for finishing a project in (3). Additionally, it can be argued that the use of semi-modal in (2) is motivated by habitual reference as conveyed with the adverb *usual*. By contrast, *has got to* in (4) expresses logical necessity in terms of the speaker's belief about a specific situation. A similar reading is found in (8), in this case with *has to*. The contracted form *'s got to* in (5) is related to external conditions such as regulations about working hours which, again, refer to a habitual action. Next, unlike the common view (see Perkins 1983: 62), in both (6) and (7) *need to* appears to be linked instead to objective circumstances, such as the need to spend more money because of the economic situation in Madras or an externally imposed necessity to understand English in school, respectively.

The last two instances (9) and (10) are interesting in two respects. In (9) there is, first, alternation between *have to* and its contracted form *'ve to*. Secondly, the objective source for buying tickets (because they were not sent to the speaker) co-occurs with the speaker's subjective belief about the presence of other partygoers (i.e. *they*) at the same event. The subjectivity expressed through the speaker in the next sequence confers epistemic reading to the utterance. Finally, in (10) the speaker seems to be unaware of a possible semantic distinction between the modal and semi-modal. While *have to* is considered to express objectivity or an appeal to reason, it is sometimes interpreted as an indicator of weakened obligation, whereas *must* would signal an "illustrative detail" (Westney 1995: 96, 106). From this example, however, it is not very clear which of the two parties (*Mr. Pepper* or *the judge*) impose the obligation or necessity. Ultimately, due to the spontaneous character of these utterances alternations can be also seen as a sign of self-correction.

In spite of some dis-fluencies and style shifts (e.g. informal use of contracted forms *'s got to* in (5) and *'ve to* in (9)) which are typical for non-monitored on-line

speech production, alternation of modal items is a common phenomenon in present-day English. The alternations in these examples reflect not only the variable choice of either a modal or a semi-modal, but also the semantic distinctions (root vs. epistemic) which these markers possess. Relevant to the discussion is the semantic nature of root *must*, *have to*, *(have) got to* and *need (to)*. It has been often argued that the modal *must* correlates with subjective necessity, whereas *have to* only relates to externally/ objectively imposed necessity or that it conveys neutrality (see Brinton 1991; Coates 1983; Groefsema 1995; Palmer 1990 [1979]; Tregidgo 1982). Conversely, the opinions on *(have) got to* vary to a great extent, with some describing it as more impersonal than *must*, and thus closer to *have to* (see Quirk et al. 1985) or as a marker of personal obligation (see Myhill 1995; 1996). Likewise, *need to* is seen as a rapidly spreading candidate to replace *have to* (see Nokkonen 2006; Smith 2003).

In view of the various positions in the literature so far, the present analysis approaches ongoing modal change as related to fluctuations in the contrast between subjective vs. objective root necessity found in New Englishes. Of interest is whether the assumed correlation between subjectively/objectively imposed root readings of these four markers with their frequency of occurrence holds in the three varieties.

1.3 Aims and scope

The general aim of this study is to integrate diachronic and synchronic evidence of ongoing semantic change in both native and non-native varieties of English. The idea embraced throughout this study is that the evolution of Standard English is not monolithic and that innovation is related both to language use and to language structure in the form of regional and stylistic patterns.[5] Thus, the four modal items under study are considered to be part of the common inherited EModE inventory sharing the same semantic domain, but which have developed at differential speeds in New Englishes. The four markers of 'strong' root/non-epistemic obligation and epistemic necessity were chosen because, semantically, their behaviour is distinct from 'weak' markers such

[5] On the mutual relationship between change in grammar and language use, see Traugott & Dasher (2002: 6): "Our assumption is that structural and communicative aspects of language shape the form of grammar […]." For an utterance-based evolutionary perspective on language change, see Croft (2000).

as *should*, *ought to* or *supposed to* (Bybee et al. 1994: 186; Smith 2003: 242). Additionally regional preferences in the dialogues from ICE are interpreted in regards to linguistic retention or innovation, namely whether traditional forms from earlier stages (*must*) prevail in some varieties or whether newer forms (*have to, (have) got to, need to*) are rapidly spreading or specializing in others (see Tagliamonte & Smith 2006).

Following the principles from corpus linguistics the present study integrates regional and stylistic patterns to uncover ongoing developments in the system. In other words, grammatical variation in the use of modals of obligation/necessity will be approached through regional, social or stylistic fluctuations (Aitchison 2001[1991]: 39-42). Accordingly, the four expressions will be seen as representing different stages or layers in the ongoing process of semantic change, which may result in functional specialization (Hopper 1991) of modal obligation/necessity. It is the way in which objective and subjective modality, respectively, correlate with processes of incipient language change which will form the focus of this study. In view of such tendencies, the low distribution of epistemic readings in the database (see Chapter 6) may reflect different dynamics (see Dollinger 2008; Tagliamonte & D'Arcy 2007) in the three new Englishes. It is useful to take note of the historical changes involving creoles or of the evidence from first language acquisition according to which ongoing change – e.g. extension towards epistemic meanings – is predictable if root meanings appear as already grammaticalized elements at a certain point in time (Shepherd 1982: 322).[6]

A further issue concerns the way in which synchronic behaviour of linguistic items is extrapolated from the successive diachronic stages of their development. Due to the nature of the data sample under analysis, a synchronic view of such changes in the apparent-time paradigm[7] will prevail, rather than a diachronic one. Nevertheless, diachronic evidence will be drawn from the available literature to support the findings from the data. A further point calls for explanation; as it will become clear, the present study does not refer to a true "apparent time sociolinguistics study" as it lacks relevant speaker information, e.g. age (for details see section 4.7.1), which would allow for comparison across generations at a specific point in time. Relevant to the discussion is

[6] Such a trajectory of development from root to epistemic modality appears to be common not only for the case of English (Traugott 1989) but also cross-linguistically (Bybee et al. 1994).

[7] As a common approach in sociolinguistics, apparent-time evidence of synchronic data paved the way for the study of variation and change in progress (see Bailey 2002; Labov 1994: 43-72; McMahon 1994: 239-243).

that diachronic developments of modal necessity in corpus data from New Englishes will be necessarily extrapolated from synchronic variation and will be assessed in a way that is similar to the apparent time approach.Against the diachronic and synchronic progression on the modal cline (e.g. *must* > *have to* > *(have) got to* > *gotta*; and *need* > *need to*), two possible scenarios emerge from the present data: (i) the linear or **monodimensional** model primarily suits the alleged notion of colonial lag, namely that one variety always appears to be more advanced and that the newer varieties are lagging behind (see section 2.2). Such a view corroborates the assumption that AmE is the most dynamic variety to date, followed by BrE as more conservative, but eventually catching up. Naturally, IrE as a newer native variety (although in the face of the historical circumstances a language-shift variety), is assumed to be closest to BrE.[8] Due to geographical proximity and the loosening of its historical British influence, JamE should be closer to AmE (Mair 2009b). Last but not least, IndE as a second-language variety is identified as being in the least favoured position in this picture, namely lagging behind as the most conservative of all the varieties under examination (Collins 2009b; Mair 2009b). (ii) By contrast, a **multidimensional** model of change is superior to the previous one, as it incorporates both the linear diachronic development as well as possible interferences from local substrates (e.g. *hafi* from JC) or from AmE which trigger globalization processes. Within such a view the ordered distribution of the three varieties in relation to BrE and AmE as hypothesised in the first scenario appears to be descriptively insufficient. Moreover, the significance of colonial lag seems to be blurred becoming "largely myths as far as the hard linguistic facts of language varieties of English are concerned" (Görlach 1987: 55). Since I wish to put contemporary English in perspective, a multilayered interpretation of the colonial lag (cf. Bauer 2002: 6) will be preferred across the whole investigation, as:

> What is more, it may be that BrE and AmE represent two extremes of a grammatical continuum, with BrE at the conservative pole and AmE at the progressive pole. Corpus-based studies include Indian, Australian and New Zealand English have shown that these national varieties are located between the two extremes in relevant respects (see, e.g., Sayder 1989, Hundt 1998a). It will therefore be a worthwhile enterprise to extend the angle to other varieties of

[8] Such patterning is confirmed, for example, in Mair's (2009a: 1116-1117) assessment of variable prepositional usage (e.g. *from*, *to* and *than* after the adjective *different*) on the World Wide Web in contemporary English.

> English spoken around the world, which can be expected to exhibit their own characteristic grammatical divergences. (Rhodenburg & Schlüter 2009: 423)

The systematic analysis incorporates three dimensions which take into account several distinctions: regional vs. stylistic diversity in formal and informal settings, and printed vs. non-printed media. The analysis focuses on the semantic (root vs. epistemic), morphosyntactic and discourse-pragmatic properties of obligation/necessity as found in spontaneous speech production.

From a usage-based functional perspective discourse frequency acts as potential motivation for the ongoing grammaticalization of items (Bybee 2003). Assuming that conventionalization of meaning is mediated by the repetitive use of elements involved in grammaticalization processes (Hoffmann 2004: 172), emphasis will be placed on the role of relative frequency in the use of the four modal items (Krug 2000: 251).[9] The in-depth analysis includes specific variants with high stylistic effect such as contracted forms '*ve to, '*s *got to*/ '*ve got to* study. Additionally the occurrence of such forms in spoken discourse might highlight the relationship between so-called *string frequency* and univerbation (see Krug 1998b; 2003).

Next, due attention will be addressed to the way in which both the public and private domain in these three varieties reflect linguistic ecologies. For example, the different socio-historical contexts in the transmission and acquisition of English as well as the influence of various other local languages can explain the behaviour of these expressions in these new standard varieties. Another observation concerns the fact that the expression of obligation/necessity varies culturally in terms of the assigned authority exerted over other individuals or groups to fulfil an action, and of their source (religious, legal or moral) which varies in each society (Narrog 2010: 408-409). Although these assumptions are usually based on cross-linguistic evidence, they enrich the overall interpretation of linguistic variation in data from JamE, IndE and IrE as divergence from or convergence to/integration with the accepted standard international norms of English.

[9] While there is general agreement on the importance of frequency of occurrence as a variable in the analysis of grammaticalization (Bybee 2007; Bybee & Hopper 2001) it should not be considered as absolute, and, therefore, all its aspects should be differentiated (Hoffmann 2004: 189-191). See also Teubert (2005: 5) for a related discussion on the role of frequency. For a critical empirical assessment of the contrastive effects of high discourse frequency, see Krug (2003).

To be precise, the specific goal of the present endeavour is to go beyond a simple description of structural patterns of strong obligation/necessity and capture the various dynamics of ongoing change in each of the three varieties (Lindquist & Christian Mair 2004: X). The main research questions which will guide my study are:

- Can a reorganisation of the obligation/necessity system in apparent-time in new standard varieties of English providing the focus of this study, namely Jamaican English, Indian English and Irish English be identified? If so, how is this related to ongoing change in this area of grammar in British English and American English?
- What are the discourse motivations for the variability and change observed and what is the relationship between the morphosyntactic environment of the verb and its modal meaning?
- To what extent is the interface between semantic (root vs. epistemic), pragmatic (e.g. source of necessity) and stylistic features (formal vs. informal) a valid indicator in mapping ongoing change in modality in other English-speaking communities?
- Does the use of strong obligation and necessity in the present data display innovations (as opposed to retention or 'colonial lag')? Are these differences significant?

In a nutshell, this study seeks to investigate (a) whether educated standard JamE, standard IndE and standard IrE still largely follow British norms of usage rather than North American or independent ones, and (b) to what extent style (formal vs. informal) is a relevant determinant in processes of ongoing change. From a methodological perspective, it is hoped that the present approach meets the desiderata of an ambitious project such as the *International Corpus of English*: to provide systematic grammatical description and comparison across the various genres and registers of the major national and second-language varieties of English.

2 The dynamics of the New Englishes

2.1 Introduction

This chapter addresses the main terminological and methodological issues for defining the New Englishes as a subfield of linguistic investigation. In this study, the New Englishes are considered as those varieties which emerged as the result of complex processes during colonial history, which have probably intensified during the recent period of political and cultural decolonisation (see Schneider 2007). The different standard and non-standard or non-native varieties of English have often been defined as an extension of colonial history, especially as regards their acceptance of either British or American norms. In the following, the sections 2.2 and 2.3 present the most common issues as regards the terminology and research directions on postcolonial varieties of English.

2.2 Terminology and categorization

The label 'New Englishes' as introduced by Pride (1982) and Platt et al. (1984) refers to the way varieties in the former British colonies (e.g. Indian English, Singapore English, African Englishes, and Caribbean English) are perceived in linguistic investigation as distinct from British English as the 'parent'-variety and the standard norm in various speech communities. Essentially, Platt et al. (1984: 2-3) stress that New Englishes have developed through the education system in a non-native speech community, with the new variety serving as an 'associate' language for communication purposes in the administration, media or in the family and with friends. The term thus focuses on the diversity of linguistic developments other than those found in native societies.[10] In contrast, World Englishes, as introduced in Kachru (1988), is a more general term, and refers to a conceptual definition according to the functions performed by English in different countries world-wide: as a native language (ENL), a second language (ESL)

[10] The opponents of this term consider it as too narrow and inconsistent, arguing that some of the non-native varieties, for example those in South Asia are historically older than the newly formed native varieties with strong British legacy, as for example New Zealand English (Hickey 2004b: 504). See also McArthur (2003).

and a foreign language (EFL). In addition, Görlach (1991: 12) distinguishes speech communities where English is used as a second dialect (ESD), such as is the case of Scotland or the Anglophone Caribbean region.

In Kachru's (1988: 5) widely known representation of the World Englishes as three concentric circles – the *inner circle*, the *outer circle* and the *expanding circle* – the genetic element prevails in the form of the 'norm-provider' (BrE or AmE) within the inner circle, whereas the 'norm-developing' second-language countries belong to the outer circle (e.g. India, Ghana, Nigeria). The expanding circle (e.g. China, Japan, Israel, and Korea) does not share the same status with the 'norm-developing' varieties being especially performance oriented, and is therefore considered as 'norm-dependent'. Unlike Platt et al. (1984: 7-9),[11] it is striking that Kachru does not include in his model for example the post-creole Caribbean English varieties, which shows the difficulty to draw neat boundaries. Although in the present context the term World Englishes seems to be more appropriate (given that a significant number of speakers acquire near-native competence of English in communities where there is no documentation of a colonial past), none of the two terms is considered precise, shedding light on inherent inconsistencies.

For the study on variation patterns within modality across three different varieties, I will take into consideration Kachru's (1988) classification referring to the status of English in the formerly British colonies as belonging either to the inner or to the outer circle. However, in order to avoid overgeneralization and in the light of the recent efforts towards standardization of these varieties I have chosen to refer to New Englishes throughout the present study. Nevertheless, the two terms are interchangeable to some extent (see also Figure 4.3). Two definitions of 'Standard English' seem relevant in this context:

> Standard English is of course that variety, or set of closely related varieties, which enjoys the highest social prestige. It serves as a reference system and target norm in formal situations, in the language used by people taking on a public persona (including, for example, anchorpersons in the news media), and as a model in the teaching of English worldwide. (Kortmann & Schneider 2004: 1)

[11] It should be, however, pointed out that Platt *et al*. (1984: 7-8) do not include in their classification of 'New Englishes' the several creoles and pidgins either. Instead, they restrict their classification to the acrolect (e.g. standard Jamaican English) and some of the mesolect varieties which are part of the post-creole continuum.

> Although it is more difficult to apply the notion of a standard to spoken English, a similar approach can be used: we define standard spoken English as including grammatical characteristics shared widely across dialects, excluding those variants restricted to local or limited social/regional varieties. (Biber et al. 1999: 18)

While Kortmann & Schneider (2004) restrict the written standard as a model for the educated users of a variety, Biber et al. (1999) broaden the notion of 'standard spoken English' to those features shared across dialects of the same variety. Although the two definitions seem to reflect different perspectives of what a standard means, they do not exclude each other. On the contrary, internationally shared grammatical features across these varieties seem to coexist with local and/or regional variants, which have become part of the standard norm.

Other terms which have been used to describe the newer varieties of English are, for example, the 'English Language Complex' (McArthur 2003: 56) or 'Postcolonial Englishes' (Schneider 2007). For Schneider (2003; 2007: 4), the diverse multilingual and socio-historical backgrounds are essential criteria for a systematic understanding of linguistic variation within these new standard varieties. A similar approach is adopted in Mesthrie (2006: 382) as well as Mesthrie & Bhatt (2008: 12) who also popularize the notion of ELC, however extrapolating the evolution of New Englishes to a multilingual history of English, which is essentially linked to the linguistic processes that took place in the British Isles before the settler embarked on the colonial expansion. These authors introduce a typology borrowing McArthur's (2003: 56) term of the 'English Language Complex', a cover-term which integrates the new varieties in the diachronic phases undergone by insular English. Much the same with the contact situations during the historical development of English – usually defined into the stages OE, ME, EModE and ModE – the diffusion of these new varieties led to contact features from at least two directions: the dialects of the settlers and the indigenous communities with their local idioms.

Schneider's (2007) monograph on *Postcolonial Englishes: varieties around the world* approaches these varieties as a direct result of colonial and postcolonial history arguing very much in the spirit of Mufwene's (2001) theory of the 'ecology' of

language.[12] Very similar to the biological sense of the term, the ecology of language refers to the selective impact of ecological conditions, such as internal and external factors along the evolution of a linguistic species (Mufwene 2001: 22, 30). Following this line of argumentation, Schneider (2007) proposes a new framework called the 'Dynamic Model' of the evolution of postcolonial varieties which incorporates five distinct stages – 'Foundation', 'Exonormative stabilization', 'Nativization', 'Endonormative stabilization', and 'Differentiation'. The five stages cover a whole range of factors such as those relating to social identity construction and reconstruction, as well as to the new sociolinguistic conditions or the linguistic effects (Schneider 2007: 6).[13] The model offers a comprehensible picture of the postcolonial varieties by avoiding the delimitation of these according to native or second-language varieties and by placing more emphasis on linguistic features. The comparison of regionally divergent varieties is two-fold, as it involves both the colonizers and the colonized. Despite the limitations of the model – as it does not offer conclusive results for all the stages (Schneider 2007: 29)[14] – the descriptive and analytic tools emphasise the emergence of postcolonial Englishes as a systematic and robust process of a sequence of diachronic events marked in linguistic structures.

2.3 The New Englishes in linguistic research

Some approaches on the New English have attributed the alleged conservative features as well as the recent independent innovations in the pronunciation and lexis to the so-

[12] Mufwene (2001: 1) argues that creoles have developed on the basis of the same restructuring processes as noncreole languages stressing the importance of contact and external factors or 'ecologies' which influence language vitality. The restructuring process of the contact between English and the indigenous languages is best illustrated as a 'feature pool' from which speakers select the elements that suit their own way of expression (Mufwene 2001: 4-6).

[13] Each evolutionary stage is presented according to four parameters: extralinguistic (sociopolitical) background; identity constructions; sociolinguistic conditions (contact settings and participant's use of specific varieties; norm orientations and attitudes); and typical linguistic consequences (structural changes on the level of lexis, pronunciation, and grammar) (Schneider 2007: 33-55).

[14] Most recently, Leitner (2013) criticized Schneider's Dynamic Model and other similar models on the ground that they do not consider "language contact as reciprocal and multi-layered in multilingual habitats"; and pointed out the need to study the postcolonial varieties of English precisely against such background.

called ‘colonial lag’. Initially introduced by Marckwardt (1980),[15] this term was applied especially in comparative analyses of British and American English (Görlach 1987; Hundt 2009; Kytö 1991). In a wider sense, the term suggests that, in the long run, the new variety will eventually ‘catch up’ with the homeland variety (Bauer 2002: 5), e.g. native varieties spoken in the United States or the Southern Hemisphere (e.g. Australia or New Zealand). However, the term can be problematic particularly when we analyse differential changes in varieties of English, as it suggests isolation from the metropolitan developments. Instead, ‘extraterritorial conservatism’ as introduced in Hundt (2009: 32) appears to be a more appropriate choice as it “includes both colonial and post-colonial language use and avoids the negative implications of ‘lag’.”

For Görlach (1987: 45) the innovations are more prominent in the speech of colonial/extraterritorial varieties than conservative features. Görlach (1995: 41-42) subsequently emphasises the innovative character of these varieties, and suggests that the discussion, whether New Englishes (including postcolonial varieties) are more conservative or innovative, must be examined in the light of language contact. A noteworthy case in this respect is that of ‘convergence-to-substratum’ with modal *must* in SingE (Bao 2010), which has developed only deontic/root meaning due to the contact ecology with the substratum. Conversely, Hundt (2009: 15) suggests that such patterns of assumed ‘lag’ are rather the result of “regressive divergence”. With respect to syntactic lags, Görlach (1987: 55) sees a possibility of identifying conservative features especially in ESL varieties, which due to “linguistic insecurity” stick to prescriptive grammatical correctness, and concludes:

> Although syntactic lags are, then, possible under certain circumstances of isolation or different educational policies as regards prescriptive ‘correctness’, there has been less of this in the overseas history of English than in other fields, and even less is likely to evolve in a world characterized by increasing communication (Görlach 1987: 55).

The question whether ‘colonial lag’ is more typical in the New Englishes depends ultimately on the definition of the term itself.[16] Accordingly, Görlach (1995: 44) argues

[15] It is in fact the concomitant alternation of “innovations and the unbelievable archaisms of the colonies” that struck Marckwardt (1980: 69).

[16] Hundt (2009: 34) considers both terms ‘colonial lag’ and ‘colonial innovation’ as useful in the synchronic descriptions of differential changes between, for example, AmE and BrE; however, as regards the first term it should be used with caution, particularly because – as the metaphor of ‘lag’ suggests – the relationship between the two varieties cannot be defined as a linear model of language change.

that colonial lag in ESL countries is expressed on other levels than in ENL countries, and largely depends on four factors: e.g. the sociolinguistic context after independence; the lack of a national norm; restriction of English to certain domains of activity; and influence from indigenous languages. Alternative explanations suggest that the diffusion of English was enhanced by a blend of substrate-superstrate derivations (Siemund 2005: 402), such as for example in the case of pidgins and creole-based Englishes, but also 'language-shift' varieties. Perhaps the most substantial descriptions and comparisons of structural features to date covering both phonological and grammatical information of English around the world – from spontaneous spoken non-standard varieties to contact varieties, e.g. pidgins and creoles, as well as major ESL varieties – are those included in the recently edited volumes *A Handbook of Varieties of English* (Kortmann et al. 2004).

A more challenging task is documenting quantitatively the degree of internal variation and incipient grammatical innovations in varieties world-wide, including native and non-native varieties (Mair 1998: 140; Mukherjee & Gries 2009: 28). Currently, one of the most employed methods in studying linguistic variation and incipient grammaticalization processes is that of corpus linguistics which uses principled collections of data for the analysis of naturally occurring patterns in language (Biber 2010: 159; Mair 2002b; 2004; McEnery et al. 2006: 3-12). Thus, comparability between varieties has been extended to other native and non-native English-speaking regions world-wide. For example, the compilation of the *Kolhapur Corpus of Indian English* by S.V. Shastri at Shivaji University, Kolhapur enabled comparability with the Brown-family corpora even though the two corpora contained samples from different periods. The discontinuity of the sampling period in the case of the Kolhapur Corpus (i.e. 1978) was claimed to be in fact an advantage, as structural nativization of Indian English started to be visible especially in post-Independence India (Shastri et al. 1986 quoted in Sedlatschek 2009: 38). Since then, similar corpora were compiled for further national varieties, such as the *Macquire Corpus of Australian English* and the *Wellington Corpus of Written New Zealand English*. The time lag in the compilation of these corpora as compared to LOB and Brown is of 25 years.

The growing interest in analysing other varieties of English culminated in the compilation of the *International Corpus of English* (ICE) (Greenbaum 1996) which

focuses on the regional spread of English as a native or second language. Due to the technological progress in the–last part of the twentieth century ICE includes besides written texts also spoken data from private and public conversations in all varieties of English across the world. The advantage of such a level of standardization ensures direct comparability with different regional varieties, as in the present study with ICE-India, ICE-JA and ICE-Ireland. In addition, the ICE-project attempts to cast new light on such concepts as 'educated' and 'standard' English (Greenbaum 1996: 28).

2.4 Summary

This chapter has presented some important directions in the study of New Englishes, showing the way in which the complex evolution of the New Englishes is linked with the different socio-historical circumstances in the various speech communities. This is reflected particularly in the terminology employed in various descriptive studies such as 'colonial lag', nativization etc. Moreover, it is particularly in the postcolonial era that these varieties have come to exhibit various degrees of nativization, whereas the concept of colonial lag appears to be descriptively insufficient.

In the light of the present endeavour, an analysis of ongoing changes in regional varieties of English requires a double perspective: (i) newer varieties of English have been acknowledged as undergoing a so-called process of 'structural nativization'; (ii) at the same time, this process entails differential stages of nativization, and thus distinct paths of development (see Schneider 2007).

3 The emergence of new standard Englishes in Ireland Jamaica and India

3.1 Introduction

Historically, the spread of English in Ireland, Jamaica and India is a contact-induced phenomenon initiated by English settlers involving subsequent bilingualism and language shift in the first case, and by colonial movements in the British Empire in the latter two. As for the evolution of English in the three countries, it can be said that it took place on two dimensions: it has followed the paths of nativization and indigenization becoming one of the country's official languages alongside Irish (Harris 1991: 38) – and Jamaica, where it is accompanied by the widely used basilect and mesolect forms of Jamaican Creole. Conversely, in India English has maintained its status of a co-official language in administration and education. Perhaps the fact that the language shift in Ireland started at approximately the same period (beginning with the 17th century) as English was transplanted to Jamaica or India, though with different effects on their development, makes the three varieties particularly interesting for our study.

To this day, in all these three cases the discussions concerning the recognition of English as a standard national variety are related to the establishment of a set of unified norms. The most difficult task seems to be the establishing of a distinct local norm which is neither vernacular nor an inner-circle standard, as described for 'Standard Irish English' in Harris (1991), Hickey (2007: 26-29), Kirk & Kallen (2005), and Mac Mathúna (2005); for emerging 'Standard Jamaican English' in Christie (1989; 2003), Mair (2002a), Mair & Sand (1998), Shields (1989) and Shields-Brodber (1997); or for 'Standard Indian English' in Kachru (1983) and Schneider (2007: 161-173). An extensive discussion of the various theoretical implications of these approaches would be, however, beyond the ambits of the present purpose. Therefore, a brief overview of the general tendencies is offered to the extent that these reflect the emergence of a distinct Irish (section 3.2), Jamaican (section 3.3) and Indian (section 3.4) variety as integrated into the wider notion of New Englishes.

3.2 Irish English

3.2.1 The shift from Irish to English

According to Hickey (2004a: 83-88), two distinct periods determined the spread of English in Ireland: the first one is associated with the first settlements in the Late Middle Ages, and the second one marks the shift from Irish to English from the seventeenth to the nineteenth century. The settlement of the Scots in Ulster at the beginning of the seventeenth century marked a radical change in the linguistic landscape of Ireland. Thus, besides Irish Gaelic, English began to interact with Scots, the latter leaving a strong imprint on the current northern Irish dialect boundary. Another significant factor which accelerated the shift from Irish to English was the establishing of National School in 1831 where Standard English was taught (Filppula 1993: 19; Joyce 1910: 8).

As the shift of the Irish speakers to English happened rather rapidly, taking place from the seventeenth through to the nineteenth century, the literature talks about a situation of imperfect adult second language learning. However, Hickey (2007: 121) comments that during the shift from Irish to English an official situation of bilingualism in the sense of a functional diaglossic distribution of the two languages did not exist in this case. Eventually, the use of Irish was gradually eliminated, and currently the Irish-speaking population lives in small areas on the west coast in Donegal, Connemara, Gaeltacht, Dunquin. By contrast, the northern counties are dominated to various degrees by Scots, as for example in Antrim, Derry, Tyrone, Armagh, Monagham, Belfast.[17]

From a World Englishes perspective it is important to mention the contribution of the Irish English inventory in the process of language transportation in overseas British colonies: e.g. in the USA, Canada, New Zeeland, Australia, the Caribbean and the South African Indian English varieties (Hickey 2007; 2004d). This phenomenon has extended the possibilities of comparative and cross-variety investigations on the legacy of Irish features in global English.

[17] Cf. Hickey (2007: 442, Map A6.5 in Appendix 6).

3.2.2 Standard Irish English as a linguistic research area

Linguistically, the emergence of English in Ireland is defined by the rural-urban distinction and stretches across the political boundary between south and north. By extrapolation, vernacular Irish English is seen as "the familiar pattern of a continuum of varieties ranging from least to most standard-like" (Harris 1991: 39). Similarly, in the conclusion of his study on the uniformity of Irish dialects which have influenced the emergence of English in Ireland, Filppula (1991: 59; 1994) supports the idea of *continua* in grammatical variation within IrE.[18]

Different terminology has been used to describe the linguistic and historical role of the Irish-English contact situation. The most common terms, i.e. *Anglo-Irish*, *Hiberno-English* or *Irish English*, are often considered as near synonyms, but at the same time are the source of various confusions (Hickey 2007: 3). While *Anglo-Irish* refers primarily to the literary written variety of English, *Hiberno-English* has been widely used to describe the various vernaculars of English spoken in Ireland. Studies on *Hiberno-English* use the term in a wider sense focussing especially on its origins as a contact vernacular (Filppula 1990; 1993).[19] By contrast, *Irish English* (IrE) is more appropriate for the context in which the linguistic boundary between north and south is not uniform (2007: 4-5). It is this term, too, which will be used for the present purpose.

The Irish influence or the continuity of older English elements is commonly reflected in language-internal variation. Three concepts largely employed in language-contact are the focus of most of the studies on IrE: the substratum and superstratum perspective, and adding to these the 'universalist' dimension. A systematic picture of the general trends in the study of the emergence of the Irish language situation is summarised in Filppula (1993). By and large, it has often been argued that retention and/or transfer is the underlying cause for the non-standard tendencies which Irish English displays in phonology or syntactic structures: e.g. the *after*-perfect, *A new fella* ***is after taking*** *over uhm one of the pubs at home* (ICE-Ireland.S1A-046.txt); the medial object perfect, *They have a local pub bought there* (Siemund 2005: 294); cleft-constructions, […] *and when they are together, 'tis* ***their own language they speak***

[18] In his study, Filppula (1991) uses the label Hiberno-English to designate the English variety spoken in Ireland.
[19] For further details, the reader is referred to Kallen (1997: 20) and Hickey (2007: 4).

***together**, German and the French*, and topicalization, e.g. […] you'd make sixty baskets in an hour […] and that was about a load. *A dry load of weed it was* (Filppula 1999: 242).

According to these approaches, historical-comparative methods were extensively employed in reconstructing, where applicable, the earlier non-standard morphosyntactic features both from Irish and English dialects, which would offer a better understanding of the contact situation (Siemund 2005: 283f). Especially substrate elements which have survived this transition have often been considered to be the source of informal features in Irish English (Mesthrie 2006; Siemund 2005: 288). As different as these approaches are, scholars generally agree that none of these sources can exclusively explain the nature of IrE. Instead, there is common agreement that both substrate as well as superstrate elements have contributed – though to different degrees – to the evolution of the new variety (Filppula 1993).

In addition to internal change documented during the process of language shift, Hickey (2007: 309) argues that the *superimposition* of speakers conveyed more standard forms of English (e.g. in pronunciation), which eventually led to *supraregionalization* and layering. This particular type of language change is different from dialect levelling and has been studied mainly in pronunciation (Hickey 2007: 309, 311f). Moreover, Hickey (2005) argues that parallel strategies must have occurred in the emergence of other English standard varieties in the world, too.

Although the particular shifting process in Ireland raises questions about the Celtic influence in general, Kirk & Kallen (2005: 91) point out that interlanguage phenomena as the result of restructuring or informal learning are not necessarily integrated in the standard variety. In their exploration of the recently completed component of ICE-Ireland, Kirk & Kallen (2005: 108-109) conclude that discourse frequency of Irish-based dialect lexicon does not seem to be a prominent feature in standard Irish English, but that Celticity operates rather at the lexico-grammatical level.

On the other hand, other linguists have proposed that all these features are nothing else but evidence of the independent evolution of IrE (see Siemund 2005). Notable is the extension to epistemic uses of modal *must* in the negative form which can be considered as a generalisation of Irish speakers based on positive use (Hickey 2007: 282).

3.3 Jamaican English

3.3.1 The creole-continuum model and the standard norm

Variation is a recurrent aspect extensively studied in extreme cases of mixed languages such as pidgins and creoles, with JC as a prominent case. A significant impact in the study of such mixed languages has had the 'continuum'- hypothesis, which shifted the attention from the problem of the origins of these languages on the mechanisms of change in post-creole varieties, as formulated initially in the 'life-cycle' theory (DeCamp 1971; Holm 1988: 52; Rickford 1987: 32).[20] According to this theory, creoles emerged out of an initial stage of pidginization in co-existence with its European-based language, thus emphasising the importance of continuous restructuring of substrate elements.[21] From a variationist linguistics perspective the concept of creole-continuum refers to a situation in which the creole variety coexists with the standard language, in this case English, displaying systematic and ordered transitions from basilect to acrolect as regulated by implicational scaling of features (Patrick 1999). Studies on Anglophone Caribbean creoles have used these theoretical views to foster creolist research as a linguistic discipline. Of these, the Jamaican contact situation is by far one of the most well studied, as it has often been considered to be a "canonical example" (Patrick 2004: 407) for the whole region.

From a sociolinguistic perspective, the foundation of JC is related to the emergence of a distinct social and ethnic identity born out of the interaction between different languages of West African slaves (e.g. Akan, Twi and Kwa group) and British colonisers in the eighteenth century (Patrick 2004: 408). In contrast, Standard Jamaican English largely originates in a blend of various EModE regional and non-standard dialects (e.g. Scottish and Irish) which were adapted to the local realities of the time

[20] The various theoretical approaches on continuum models have identified two main directions of research: *discreteness* and *unidimensionality*. The various continuum models were concerned to identify: (a) whether creoles exhibit discreteness of categories or a gradation of various intermediate sequences of 'lects' between one pole to the other; (b) as well as, whether variation is linearly ordered on one dimension representing either 'creoleness' or 'standardness' (DeCamp 1971; Rickford 1987). Conversely, the structuralist tradition sharply separated the two linguistic systems, i.e. the local creole variety and the standard norm, as each of these were considered to function according to a predefined set of rules or patterns (Bailey 1971: 341).

[21] Subsequent scholars have challenged this theory, as for example Mufwene (2001; 2006), who argues that there is no genetic relationship between pidgins and creoles and that these emerged due to independent ecological conditions.

(Christie 2003: 8-9; Lalla & D'Costa 1990: 6). Moreover, the fact that the first British settlers had little education or that some of them were even illiterate, contributed essentially to the spread of these non-standard dialects. Adding to these, the African input in JC represents an important substrate element contrasting with Standard English norms, which can be observed particularly in the grammatical structure of the basilect.

A steady *decreolization* process with urbanisation and social reforms started in the post-emancipation period in Jamaica, creating opportunities for schooling among the poor, as well as for social mobility (Christie 2003: 11; Patrick 1999: 26; Schneider 2007: 234). However, Standard English remained for a long time the language of a small group representing the upper class, which followed as target the British norms. In linguistic terms, decreolization refers to a tendency of replacing the most non-standard features with those which are closer to the norms of Standard English, generating a continuum of various intermediate options widely known as: basilect - mesolect - acrolect (Holm 1988: 9). Being exposed to Standard English only through the education system, the competence of the majority has often been characterized as deriving from a range of various independent features that are not exclusively basilect but neither acrolect. This has lead to a situation in which the mesolect – the intermediate variety of variable but systematic ordering of English forms and rules, and which are also typical for the basilect – has currently emerged as the most important variety among Jamaicans (Patrick 2004: 409-410). In other words, an extreme polarization in terms of random variation or interference between the invariant grammars of these two ends, such as often found in earlier linguistic descriptions, fails to account for the complexity of the speech situation in present-day Jamaica. This, in turn, brings us closer to the debates on DeCamp's (1971) concept of a 'post-creole continuum' which attempted an overarching explanation of the finely graded span of various intermediate structural features in creole speech communities.

At the heart of this concept are so-called post-creole speech communities "in which a creole is in the process of merging with a standard" (DeCamp 1971: 349, 351). DeCamp redefines the concept of synchronic linguistic variation focussing especially on Jamaica, where he identifies high degree of variability. The most significant observation is the fact that variability between the lowest and the highest form of speech represents

an unbroken, continuous line of varieties and that few speakers use the 'pure' basilect. This in turn puts in perspective the nature of the standard norm:

> The 'standard' is not standard British, as many Jamaicans claim; rather it is an evolving standard Jamaican (or perhaps standard West Indian) English which is mutually intelligible with, but undeniably different from, standard British. (DeCamp 1971: 350)

However, in post-creole societies the emergence of the two extreme poles is determined mainly socially, as the opportunities for the population to learn the standard are unequal from the very outset (see Christie 2003; Patrick 1999; 2004; Sebba 1993). Since access to the standard language is often limited to certain groups of the population, the result is that of diffusion, namely that non-standard features are not necessarily eliminated, but are likely to transcend the standard norm (Holm 1988: 53). In view of such transfer situation, linguistic variability is intrinsically linked to the social stratification in Jamaica (Patrick 2004: 408).

Thus, the idealised concept of a linear continuum of intermediate varieties which link the two extreme poles is an adequate tool to explain levelling only if one restricts variation exclusively to linguistic features. Christie's (2003: 33) examples below illustrate such a situation in the verb phrase, where minimal structural differences can be identified:

Creole:	Wi (d)a go
	Wi goin
	Wi is goin
	Wi is going
English:	We are going

As a rough generalisation, apart from various phonological, morphological or semantic influences coming from the substrate, superstrate or adstrate, syntax (e.g. the verb phrase lacking inflections or the noun phrase) appears to be a prominent marker of any creole-continuum (Holm 1994: 144-215).

Within the confines of the present study, the continuum-hypothesis is relevant in so far as JC has a very well developed modality system, and which due to the shared formal similarities (Bailey 1966) may interfere with the acrolect in spoken interaction. However, Christie (1991: 223) notes that the JC modality system as a category is overall different from what we know from Standard English. The various aspects of the JC as

well as JamE modality system will be presented at a more appropriate place of this study (see 4.4.3). Since the various descriptions of the Jamaican post-creole continuum point towards high degree of variation within and between varieties, it seems correct to take into account the possible creole influences in levelling of modal obligation and necessity in Standard Jamaican English.

3.3.2 English as an emerging standard in contemporary Jamaica

Traditionally influenced by the British norms as inherited from the various settlers during the colonial period and as propagated through the educational system Standard English in Jamaica was considered a superior model. However, the standard norm sharply contrasted with the more popular creole during colonial times and more so in the transition from the pre- to post-independence period starting with the year 1962 since when a Jamaican local/regional standard has steadily emerged (Shields-Brodber 1997: 58; Shields 1989: 43).

Currently, JC is by far the dominant and preferred variety in informal interaction, which in post-Independence Jamaica has extended to public contexts of speech (Christie 2003: 2), challenging the norm traditionally assigned as British (Shields-Brodber 1997: 57). The widespread use of basilect and mesolect forms in public discourse reflects a common device among speakers to indicate ambivalent attitude to the social status of creole as opposed to the standard (Christie 2003: 5). As Mair (2002a: 33) observes, these forms are often employed "to construct a public persona/identity, or to make a language-political statement […]". In the light of Schneider's (2007: 26) 'Dynamic Model', we could argue then that the linguistic variability as we find it in the Jamaican context is a symbolic manifestation for the construction and reconstruction of social identities.

Recently, the status of English in Jamaica has gained more attention, being acknowledged by scholars as an emerging new standard, which has loosened ties with the British standard (see Christie 1989; Mair 2002a; Sand 1999). The impact of literacy on the users is particularly salient as it generates competing models of standards as target (Shields 1989). Language debates have occupied a significant part of the

Jamaican reality in the past forty years, particularly in relation with Patois, the basilect creole, and which were paralleled by efforts to standardize an emergent variety. Attempts to redefine the notion of a standard started only in post-independent Jamaica affecting overall the attitudes about language (Mair 2002a; Shields 1989).

Perhaps more than in other speech communities, the Jamaican linguistic situation is reflected also demographically in the rural-urban divide, where Patois is mainly confined to rural areas contrasting with the mesolectal and acrolectal varieties in urban centres (Patrick 1999: 47). In addition, there is a direct link between the use of these regional varieties and education, as mastering the standard language is often associated with the educated and the elite to which the population from rural areas did not have immediate access (Christie 2003: 2).

Despite the predominance of creole in private interaction, educated written practice in post-independence Jamaica largely conforms to the standard metropolitan norms. While JamE or Patois was for a very long period restricted to particular genres (e.g. fictional texts, cartoons), Mair (2002a: 36) has identified increasing creole elements in informal written communication using digital technology such as e-mail and web forums (see also Hinrichs 2006). It is in such text-types that a continuum-like writing practice is more salient (Mair 2002a: 56). By contrast, educated spoken usage, the acrolect variety, in informal contexts is largely dominated by upper-mesolectal speech forms with basilect forms occasionally employed as a rhetorical means to create 'anti-formal' attitudes in discourse defined in Allsopp's (1996) terms as:

> Deliberately rejecting Formalness, consciously familiar and intimate; part of a wide range from close and friendly through jocular to coarse and vulgar; any Creolized or Creole form or structure surviving or conveniently borrowed to suit context or situation. (Allsopp 1996: lvii)

The range of such strategies has been recently examined in Deuber (2009). On the basis of a corpus analysis of morphological and syntactic variation in spoken data from ICE-Jamaica, Deuber (2009) concludes that informal interaction among educated speakers of English is far more complex than previously intuited. As a graded phenomenon, so argues Deuber (2009: 47), spoken interaction in Jamaica can be better described in terms of a continuum consisting of several stylistic options ranging from informal to anti-formal features.

At the same time, the linguistic variation might be the outcome of the influence that standard English has on the local variety. In particular, North American English as a major player in the process of globalization is of significance for the Caribbean Englishes:

> The influence of North American rather than British English grew in the Caribbean area as the United States emerged as a world power at the beginning of the twentieth century. (Holm 1986: 17)

> At the same time, just as British Standard English has been influenced for some time by American English, so it is likely that local varieties within the Caribbean will influence each other and will continue to be influenced by external varieties with which they may come in contact. (Christie 1989: 245)

Moreover, as a consequence of officially dissolving the British colonial dominance in Jamaica in 1962, new linguistic phenomena have started to enter the system alongside the already accepted creole forms. As commented in Christie (2003: 20), geographic proximity has led to borrowing of specific American speech habits accessible to the population mainly through the media more than in the case of other varieties.

Generally, it is argued that most of the differences between standard English and standard Jamaican English are to be found in pronunciation or lexis (Christie 2003: 14). This notwithstanding, syntax is the place in which many non-standard features have become part of the mainstream linguistic behaviour: e.g. on subject-verb concord (*Men like Blookfiels is particularly important here*), see Christie (1989: 249) and Jantos (2009); on the lack of do-support and inversion in wh-questions as a prototypical example of informal usage (*Which school you went to?*) or lack of subject inversion (*So what you're doing for summer?*), see Deuber (2009: 10-11); on patterns of verb complementation (*The MP took note the very poor conditions of the road*), see Mair (1992: 87).

Among the morphological features which are commonly assigned to creole in informal interaction of educated speakers, Deuber (2009: 46) notes the use of zero copula and unmarked past of verbs along with pronoun forms *him* and *them*. In addition to these, Mair (2009b) notes some independent innovations such as the use of *person(s)* for *people* as a sociolinguistic marker of formal style. Although 'colonial lag' or hypercorrection might be invoked as immediate explanations, the high frequency of *person(s)* might reflect instead a process of incipient grammaticalisation towards an

indefinite pronoun (Mair 2009b: 14). In the corpus-based analysis of five variables (e.g. *people*/*persons*; main-clause order in *wh*-questions; modals of obligation and necessity; contractions and quotatives) as indicative for ongoing change, Mair (2009b: 25 f) concludes that both North American influence as well as the local creole and mesolect variety seem to be the strongest competing factors in shaping Jamaican English as an emergent standard.

All in all, the present study departs from the assumption that Standard Jamaican English is largely influenced by the basi- and mesolectal creole variety with which it is linked through a continuum of ordered transitions.

3.4 Indian English

3.4.1 The spread of English in India

According to Schneider (2007: 161-173), the spread of English in India has undergone three out of the five diachronic phases described in the 'Dynamic Model', even if present-day India seems to be entering into the early stages of 'endonormative stabilization'. Similar distinctions of three historical periods in the evolution of English in India have been proposed by Mehrotra (1998) and Mukherjee (2007).

Schneider (2007: 163) sees phase one (1600-1757) as a rather long period in which first signs of bilingualism appeared, but no other linguistic influences were visible yet. Despite the long period of accommodation between the settler group and the indigenous group, nor any signs of a new identity construction could be identified at this first stage (Mukherjee 2007: 164).

It was the second half of the eighteenth century, more exactly in 1784 with the India Act, which marked an important transition from the initial trading activities of the British Crown to a powerful authority over the whole East Indian Company (Mukherjee 2007: 164; Schneider 2007: 163).[22] As a natural consequence, English started to be used more often in the administration as well. From the perspective of the 'Dynamic Model', 'exonormative stabilization' was already initiated. The British authority grew rapidly

[22] However, from a historical perspective, the beginning of the British supremacy over India is marked by the battle of Plassey in 1757, when the British defeated the last Moghul emperor of Bengal (Mukherjee 2007: 164).

gaining control over the whole subcontinent, which, eventually, became an exploitation colony.

The period 1780-1830 documents a significant increase in the establishing of English teaching schools which, eventually, contributed to the exonormative stabilization of English (Mehrotra 1998: 3; Schneider 2007: 164). Parallel to the systematic spread of English through education, bilingualism among speakers became a reality, and the first signs of a distinct Indian way of speaking English with increased lexical borrowing from Indian into English could be observed, which Sedlatschek (2009: 12) assigns as the beginnings of a 'nativization' process. Similarly, for Mehrotra (1998: 17) and Schneider (2007: 167) the nativization process continued during the twentieth century and even accelerated in post-Independence India after 1947.

At structural level, apart from the already existing practice of lexical borrowing, Mukherjee (2007: 167), Schneider (2007: 169) and Hickey (2004c: 545-546) note various phraseological (e.g. spoken and written stylistic conventions, see analysis of text types Görlach 1995), morphological (pluralization of mass nouns such as *alphabets*, *furnitures*), grammatical innovations (the use of *isn't it ?* or *no* as invariant tags; sentence structure such as in the case of *wh*-interrogative clauses without inversion; reduplication as a means for emphasis, e.g. *I bought some small small things*; use of stative verbs with the progressive, e.g. *Mohan is having two houses*), and phonological changes (monophthongization and diphthongization; speech rhythm; syllable-initial voiceless stops are unaspirated; the tendency to retroflex /t/ and /d/), which remained unknown in British English.

Although English was officially recognized in the Indian Constitution, its use among Indians was meant to be transitional until 1965, when it should have been progressively replaced by Hindi. Eventually, the Official Language Act from 1967 recognised English as a co-official language for an indefinite period alongside Hindi. At present, Hindi is still highly valued as an official language in spite of the major progresses in the functional 'range' and societal 'depth' (Kachru 1982; 1986; 1994) displayed by English.

Given these new developments of English as an intranational linking language, Mukherjee (2007: 168) sees present-day India developing into an endonormative variety. On the other hand, Mukherjee (2007: 170) rightly observes that the dynamics of

IndE are subordinated to two kinds of conflicting forces, namely between conservative and progressive forces operating at structural, functional and attitudinal level. For example, in spite of its wide-spread functional use Indian English still appears to be an exonormative variety especially in educational institutions, as too often the teaching models are biased towards promoting standard British guidelines of the formal written style. By contrast, Schneider (2007: 171) sees only early signs of an endonormative evolution arguing that such linguistic manifestations should not be overestimated. Additionally, Mukherjee (2007: 170) identifies even dialect divergence, which would be assigned to the last stage on the evolutionary cycle, whereas Schneider (2007: 172) argues that the various regional Indian Englishes reflect rather lack of a uniform national standard.[23]

All in all, currently there is evidence that IndE will continue to hold an important functional role in intranational communication. However, it seems unlikely and far-fetched at the moment to predict any long-term changes with regard to its use as a native variety among Indians.

3.4.2 Recent research in standard English in India

As a second-language South Asian variety,[24] IndE followed the standard British English norms more closely in the early stages of its transplanting on the Indian subcontinent and more loosely since its decolonisation. In spite of the ambiguous profile reflecting a range of contradictory attitudes, IndE has proved to be an established non-native second-language variety connected to the international context. Apart from English, the linguistic landscape of India consists largely of four groups of languages: Indo-Aryan (Hindi, Bengali, Marathi, Gujarati, Urdu); Tibeto-Burman (Angami, Ao, Bodo); Austro-

[23] The issue on the supposedly missing standards in IndE turns out to be more complex. According to D'Souza (1997: 94-96), the myth of falling standards is related to 'wrong yardsticks' that BrE norms should be followed as English does not belong to Indians in the sense of a native language. In response to such reasoning, D'Souza (1997: 94) emphasises the importance of accepting the local norms as standards and that more reliable descriptions of such varieties are needed.

[24] According to the 1971 census only 223,981 speakers from a population of over 900 million considered English as their mother tongue (Mehrotra 1998: 1). For further details on the demography of India the reader is referred to Kachru (1986; 1994) and Sailaja (2009).

Asiatic (Munda, Santhali, Khasi); and Dravidian (Tamil, Telugu, Malayalam, Kannada) (Sailaja 2009: 1).

Over the past few decades, features of structural nativization have been the focus of various scholarly investigations which have acknowledged its legitimacy as distinct in comparison to British English, the mother-variety. These regional features are often assumed to reflect the complex South Asian linguistic context. Mukherjee (2007: 178-181) goes even further and claims IndE to be a semiautonomous variety. Three arguments support this statement: a) the fact that IndE displays features of the "common core" of English; b) it serves as an "interference variety" because of the various local patterns influencing its structure; c) finally, many of these innovative features are rather a sign of local creativity than of interferences from the L1, which reflect its autonomy as a "norm-developing variety" (Mukherjee 2007: 179). From this perspective, the linguistic profile of IndE raises interesting questions with regard to register and stylistic variation.

Yet, the establishing of the term IndE is problematic as for many linguists and speakers alike it generally denotes 'bad English', and, thus, maintains the stereotype of low language standards in present-day India (Varma 2004: 114 quoted in Sedlatschek 2009: 26).[25] In some other cases the IndE denomination is denied precisely because main focus is placed on the diversity of regional variants (Hickey 2004c: 542). Likewise, many earlier descriptions relied on over-generalisations regarding the 'cline of proficiency'[26] which ranges from educated speakers of English to those of 'Broken English' (Kachru 1994: 509). Those who accept the term IndE consider it a "viable, vigorous variety" (Mehrotra 1998: 15) with an established status among other world Englishes. Even so, most of the linguistic descriptions on IndE so far have failed to provide a comprehensive image of the distribution of user-specific language properties as distinct or deviant from other native varieties.

Among the weaknesses in previous linguistic approaches, Sedlatschek (2009: 24-40) identifies impressionistic descriptions by means of feature lists of variety-specific structural and usage patterns (e.g. Trudgill & Hannah 2002). In this vein,

[25] See Mehrotra (1998: 6) on the 'Janus-like' profile of IndE, elsewhere also accounted for as a sign of 'linguistic schizophrenia' (Kachru 1983: 179).

[26] The 'cline of proficiency' can be determined according to three parameters: the user's proficiency in English as obtained through language acquisition and the period of instruction in the other language; the linguistic, regional, as well as the ethnic background of the speaker (Kachru 1994: 508).

Sedlatschek (2009: 28ff) draws attention to the flaws in restricting certain features as typical for educated speakers of IndE, such as the use of invariant tags with *isn't it?* (Kachru 1994: 519f), and argues that such claims are one-sided which lack a systematic assessment of register-specific variation patterns.

An exception is provided by studies which have tested these observations against large-scale machine-readable corpora of various texts of the written genre: e.g. Leitner (1991) on inter- and intravariation in the Kolhapur Corpus; Mukherjee & Hoffmann (2006) on complementation patterns of several verbs in IndE online newspapers; Mukherjee & Gries (2009) on the co-occurrence of intransitive, monotransitive and ditransitive constructions in IndE, HKE and SingE; Schneider (2004) on particle verbs in ICE-Ind; Sedlatschek (2009) on feature-specific variational profiles across registers and modes in a micro-corpus of spoken and written IndE as well as in online corpora of Indian quality newspapers; Shastri (1988) on modal auxiliaries, *if*-clauses, and verb particle-constructions in the Kolhapur Corpus of IndE; and Wilson's (2005) multivariate quantitative analysis on the behaviour of modal auxiliaries across text-types and between the national varieties in the Kolhapur Corpus of IndE compared to LOB and Brown. For example, Shastri (1988), Leitner (1991) and Wilson (2005) arrived at the conclusion that the major differences between IndE, BrE and AmE are not qualitative but quantitative in nature and that it shares similar degrees of variability as in other national varieties of English. The various syntactic and semantic aspects of variation occurring in particular with modal auxiliaries in IndE are discussed in more detail in a separate section (see 4.2.3). By contrast, the findings from studies based on larger corpora, such as Mukherjee & Hoffmann (2006) and Mukherjee & Gries (2009) point to certain lexico-grammatical differences at the level of verb complementation patterns and verb-construction associations which reflect processes of structural and collostructional nativization.

3.5 Summary

The overview of the main landmarks in the development of English in Jamaica, India and Ireland has helped integrate the current topic into the framework of New Englishes.

In the strict sense, IrE does not belong to what is called an ex-colonial variety of English. However, the fact that it resulted from a situation of language contact and shift involving English and Irish, makes it hold a special place among the other contact-induced English varieties. Starting as an L2, the English variety spoken in Ireland is exemplary for the continuous substrate elements which are identifiable as its characteristic features even after the population started to use English as a first language.

As part of the Caribbean Anglophone West Indies and a former British exploitation colony dedicated mainly to sugar production, Jamaica has historically strong ties with English. JamE has developed heavy linguistic restructuring in the form of creolization in a context of extreme social conditions. In particular, its English-based basilect variety known as ‘Patois’ has been often described as a radical case of such a linguistic development.

As the largest South-Asian English variety with regard to the number of speakers, IndE is an interesting example of the functional distribution of English in a multilingual society. Even though a former British colony, historically, IndE has not followed a continuous evolution comparable to other colonial Englishes. It is, however, deeply rooted in the society as a second-language variety.

4 Theoretical and methodological background

4.1 Introduction

The present chapter outlines the relevant theoretical foundations which underpin the general claims in research on obligation and necessity in ex-colonial Englishes. As will become clear, modal use in JamE, IndE and IrE is intertwined with empirical evidence found both in standard and non-standard varieties. The central concepts for the analysis derive largely from the current general linguistic and typological literature on modal verbs and modality (e.g. Bybee et al. 1994; Bybee & Pagliuca 1985; Jespersen 1924; Palmer 1990 [1979]; van der Auwera & Plungian 1998), from corpus-based approaches (e.g. Coates 1983; Collins 2009a; 2009b; Depraetere & Verhulst 2008; Krug 2000; Leech 2003; Smith 2003; Tagliamonte 2004; Tagliamonte & D'Arcy 2007; Tagliamonte & Smith 2006), as well as from the theoretical insights on the various mechanisms of language change (e.g. Aitchison 2001[1991]; Croft 2000; Labov 1994; McMahon 1994; Sweetser 1990; Traugott 1989; Traugott & Dasher 2002).

In the following, a review of previous corpus-based studies is outlined in section 4.2 which integrates the topic into the larger transatlantic perspective of regional and stylistic variation. Section 4.3 is dedicated to the presentation of the data sources and the methodological steps. Section 4.4 discusses the structural and sociolinguistic dimension variation and change in modal usage in the world Englishes. The relevant terminology defining linguistic modality is introduced in section 4.5 while 4.6 stresses the limitations of the present study.

4.2 Previous corpus-based research

Despite the various approaches ranging from historical studies on the development of modal auxiliaries (e.g. Denison 1993; Lightfoot 1979; Plank 1984; Visser 1969)[27] to synchronic descriptions of syntactic and semantic features (e.g. Duffley 1994;

[27] The historical development of specific modal items as well as certain semantic aspects of modality are extensively examined in: e.g. Brinton (1991), Fischer (1994), Goossens (1982; 1984; 1987; 2000), Nordlinger & Traugott (1997), Shephard (1982), Traugott (1989; 1999), Traugott & Dasher (2002) and Warner (1990; 1993).

Jacobsson 1979; 1994; Palmer 1990 [1979]; Perkins 1983; Tregidgo 1982; Westney 1995),[28] this section is confined to discussing the significance of corpus-based research on the uses of strong obligation/necessity across various registers in the major varieties of English.

Corpus-based studies on the English modality system fall largely in two categories: studies which focus on the long-term historical developments and those which analyse recent diachronic[29] and synchronic changes. The second category comprises quantitative accounts on the distribution of modals across text-types and their semantic behaviour as conditioned by various linguistic constraints across registers mainly from BrE and AmE, in some cases incorporating also variationist sociolinguistics approaches. In view of the present purpose, these two lines of research belong to the narrow focus of our interest. In spite of different research perspectives, these studies provide ample evidence that modal *must* is drastically decreasing whereas the semi-modals *have to*, *(have) got to* and *need to* are increasing at various speed in genre-specific environments in present-day English.

In the following, I shall review the most pertinent contributions to the understanding of these major shifts in present-day English. For the sake of clarity, I shall start the discussion with the diachronic and synchronic contributions on modal use in contemporary BrE and AmE (see section 4.4.1) followed by those on data from non-standard (see section 4.4.2) as well as non-native Englishes (see section 4.4.3).

[28] Although based on a relatively extended set of corpus data, Westney's (1995) study cannot be considered a systematic corpus linguistic study of modality, as regional and stylistic variation patterns between BrE and AmE are albeit briefly mentioned, and shifts in modal use are not assessed in terms of their quantitative distribution. Even so, it remains a genuine descriptive study on the modal – semi-modal alternation in present-day English. Semantic aspects of 'strong' obligation/necessity, and modality in English or cross-linguistically are treated also in edited volumes (Bybee & Fleischman 1995; Facchinetti et al. 2003; Klinge & Müller 2005), in several individual contributions (Groefsema 1995; Klinge 1993; Narrog 2010; Sweetser 1982; Tregidgo 1982; Wierzbicka 1987), and in various monographs (Nagle 1989; Nuyts 2001; Papafragou 2000; Sweetser 1990). For recent accounts on theoretical issues in modality in English also using corpus material see Salkie et al. (2009) and Tsangalidis & Facchinetti (2009).

[29] For corpus-based studies on the diachronic development of selected modals and semi-modals both in BrE and AmE, see Biber (2004), Biber et al. (1998), Gotti et al. (2002), Kytö (1991), and Loureiro-Porto (2009).

4.2.1 Modal usage in British and American English

Krug (1998a; 2000) argues that the semi-modals *have got to*, *have to* and *want to* as well as their contracted forms *gotta*, *hafta*, *wanna* display increased auxiliarisation and thus are acquiring emerging modal status in present-day English. Based on grammaticalization theory, Krug integrates diachronic and synchronic evidence of the grammatical variation of semi-modals. Focussing on the frequency of use of semi-modals in the history of English, Krug (2000: 2) claims that the set of phonologically reduced forms of modal structures (e.g. *gotta, wanna*) are indicative of change in progress. According to Krug (2000: 23) the emerging grammatical status of these expressions is closely linked to discourse-frequency in spontaneous speech where language change first occurs. Particularly in the analysis of the short-term developments of *have to* and *have got to* in fictional and press texts from the Brown corpora *incidence* proves to be a useful methodological device which measures the differential speed of change across text types (Krug 2000: 85-86).

Myhill (1995; 1996) and Jankowski (2004) provide two different perspectives on changes in the modals of obligation and necessity in AmE and BrE. Myhill (1995) focuses on the development of the functions of modals and semi-modals in AmE using a corpus with data from written-to-be-spoken drama texts in the time before and after the American Civil War (1824-1947). However, due to the limited choice of text type to one genre which lacks comparative data from BrE the study cannot be considered a true corpus linguistics analysis. Therefore, the main points of criticism are directed both to the genre selected and the narrow time frame. In fact, Myhill (1995: 204-205) is aware of this potential flaw, arguing that the selected drama texts could have been possibly influenced by post-Civil War playwriting conventions, and therefore appear stylistically more realistic. In short, changes in the function of strong obligation and necessity in AmE are determined by societal movements affecting the entire modality system. Particularly the increase of *got to* and *have to* at the expense of the decrease of *must* seem to reflect such changes. Analysing the frequency of such expressions, Myhill (1995: 165-173) tentatively concludes that the use of *got to* is an expression of "emotion/urgency", *must* is motivated by "social decorum, norms, principles, and morality", whereas *have to* is unemotional and functions as a marker of habitual

obligations. Subsequently, Myhill (1996) argues that *got to* expresses personal obligation whereas *have to* stands for objective obligations. By contrast, *must* reflects "societal expectations" (Myhill 1996: 353). Myhill (1996: 381-382) concludes that semantic change proceeds within semantic groups reflecting fluctuations in the frequency of a certain type of meaning.

Much the same focus is addressed in Jankowski (2004) on variation and change in progress in deontic modality. Unlike Myhill (1995; 1996), Jankowski (2004: 88) extends the perspective to assess the consistency of variation and change in real-time data from BrE and AmE plays sampled from 1896 to 2001. Her analysis is based on quantitative multivariate methods such as variable rules – a useful analytic tool for the correlation of factors that condition linguistic variation (Jankowski 2004: 86).[30] Even though the issue of representative data is again debatable, the choice for drama dialogue is justified from a variationist perspective as approximations of the spoken vernacular (Jankowski 2004: 89). The frequency analysis across time reveals that deontic *must* is in decline in the last century, whereas *have to* is steadily increasing particularly in AmE. According to Jankowski (2004: 95), such a trend runs counter to the well-known theory of 'colonial lag'.

To assess the consistency of such trends over time, the author correlates several language-internal factors (e.g. type of verb, reference of subject) which encode deontic modality, and which is similar to the approach in Tagliamonte (2004). The two semi-modals indicate that change is related not only to differences in frequency but more so to the way they grammatically encode deontic modality, which indicates that BrE is lagging behind AmE (Jankowski 2004: 106). The multivariate analysis reveals that *must* and *got to* are favoured with similar constraints, with the former item used in specialized contexts contrasting with its decreasing trend. The study supports Myhill's (1995) claim about the major societal changes in AmE whereas, on the other hand, it leans on Hopper's (1991) principles of grammaticalization (e.g. layering, specialization and persistence) which are connected with high discourse frequency of *have to* and *have got to* (Jankowski 2004: 108).

Comparative corpus studies on major changes in the modality system have been undertaken particularly for BrE and AmE. I should name the most significant studies as

[30] The standard methodological tool in variationist sociolinguistics is the multivariate Variable Rule analysis (see Sankoff & Labov 1979).

Biber et al. (1999), Leech (2003), Leech et al. (2009), Mair (2007), Mair & Leech (2006) and Smith (2003). Except for Biber et al. (1999), who draw their data from *The Longman Spoken and Written English Corpus* (LSWC), the other contributions infer their observations from the four matching Brown corpora. These studies are generally descriptive in character and explore short-term trends in modality mainly in written English.

Analysing the frequency rates and the distribution of modals in written data from the Brown corpora, Leech (2003) suggests that modality is "on the move". More specifically, modals are in decline, among which *must* is one such candidate, this trend being more advanced in the AmE data and consistent also across time. According to Leech (2003: 229), such tendencies reflect the spread of semi-modals which enter in competition with some of the functions of the core modals, including the cases of *going to* and *will* or *have to* and *must*. While these changes are indirectly linked to the grammaticalization of modals and semi-modals, Leech (2003: 236-237) relates these findings to two functionally-based explanations such as 'Americanization' and 'colloquialization' (Mair 1997; 1998). Another possible explanation is 'democratization' as associated with Myhill's (1995) conjecture on the development of obligation and necessity in AmE which links the shifts in the modal meanings to societal changes in the sense of avoiding overt power and authority.

In Smith (2003), ongoing grammaticalization in BrE and AmE is assessed on the basis of stylistic, syntactic, semantic and pragmatic factors which are responsible for the evolution of strong obligation/necessity. As in Leech (2003), the quantitative analysis supports the claim of the recent drop in the frequency of modals which compete with the increasing use of semi-modals both in BrE and AmE. Unlike Leech (2003), Smith (2003: 251-252) goes one step further and differentiates the frequency patterns by genre. Even though the general decline of *must* is common in both varieties, the more formal categories such as learned writing and press editorials in BrE do not seem to follow this trend. By contrast, *have to* is preferred in fiction and press genre. It is the use of *(have) got to* which reveals a striking small frequency pattern in the written data.

To understand such fluctuations, Smith (2003: 252) emphasises the recent stylistic trends in written language as observed in the differential spread of linguistic innovations across specific genres (see Hundt & Mair 1999). Within deontic necessity,

Smith (2003: 255; 260) notes that the increase of *need to* in present-day English might function as a possible competitor for *must* and *have to*.[31] The comparison appears to be pragmatically justified, in the sense that the imposed obligation meaning of *need to* (e.g. *You **need to** get a hair-cut*) is more indirect in contrast with *must*, which in its prototypical use is associated with direct speaker authority. Another significant finding represents the use of epistemic meanings as a conversational feature which seem to occur also with *have to*, however, not as common as in the case of epistemic *must* (Smith 2003: 264).

In the same vein, Mair & Leech (2006) defend the idea of the real-time study of ongoing changes as found in matching corpus material, and emphasise that significant fluctuations between BrE and AmE occur first in spontaneous discourse. Their assumption is based on the evidence found in Biber et al. (1999: 486) that the distribution of modals differs strikingly according to register. As the findings in the LSWE corpus reveal, modals are more common in conversations than in expository writing.

Mair & Leech's (2006) frequency analysis confirms the drastic decrease of *must*. Furthermore, semi-modals are most preferred particularly in spoken AmE which, according to Mair & Leech (2006: 328), is considered as the main locus of change. The study suggests that BrE – even if at a different pace – follows the AmE pattern of change in the decrease of modals.

Leech et al. (2009) is another follow-up study continuing the established tradition of comparative corpus studies. The individual chapters on modals and semi-modals draw on evidence from previous studies on the recent changes in BrE and AmE. The main contribution of this work consists in a more fine-grained corpus-based approach on language change.

Millar (2009) focuses on the recent changes in the use of modals in the TIME Magazine Corpus challenging important methodological aspects of variation in diachronic corpus data (e.g. corpus type and size or sampling variation). Re-examining the use of modals in BrE and AmE as explored in Leech (2003), Miller (2009: 193) claims that a gap of 30 years between two specific data points, as for example in the Brown corpora, is not a sufficiently reliable source to pinpoint language change.

[31] See Nokkonen (2006) on the comparison between *need to*, *must* and *have to*.

Particularly in the case of the Brown corpora, which comprise several genres as representative for the written register, variation patterns might not be gleaned that easily. In contrast with the results in Leech (2003), Millar (2009: 205) argues for a continuum in the spread of changes within newspaper data. As regards semantic shifts, Miller (2009: 209) suggests that the differences may not be only due to the variation between modals and semi-modals (e.g. transfer of epistemic meanings of *must* to *have to* or *need to*), but also among the modals themselves (e.g. tendency of *may* towards monosemy).

In addition, a number of predominantly synchronic studies are available. They all provide detailed syntactic and semantic descriptions of selected modals and semi-modals in English but do not necessarily focus on marked regional differences in usage. Instead, these point particularly to semantic aspects of modality as in Coates (1983) on the modal use in spoken and written registers of British English; Depraetere & Verhulst (2008) on the semantic differences between *must* and *have to*; Nokkonen (2006) on the semantic variation of *need to* in spoken and written corpora; and Taeymans (2004) on the alternation between *need* and *need to*, to name but a few.[32] Due to the narrow perspective, these studies provide a fine-grained analysis at micro-linguistic level functioning as a valuable resource for a wider research scope such as the reorganization of the modal system in New Englishes.

Coates (1983) remains a classic systematic corpus-based investigation which establishes semantic and syntactic criteria of modality in present-day English in data from the SEU and LOB. Based on the theory of 'fuzzy sets' in Zadeh (1972), Coates' (1983) treatment of modality in terms of indeterminacy helps to distinguish elements expressing either central or peripheral modal meanings by assigning their membership as a gradient. However, according to Palmer (1990 [1979]: 22), the approach in terms of indeterminacy is not convincing, arguing that it cannot be placed on the same par with the notion of ambiguity. In spite of the very detailed semantic examinations of modal markers, the author does not attempt any far-reaching account of ongoing change in the variability between modals and semi-modals. After all, this was not included as a research question, as the primary aim of the study is "to interpret data, not to impose some neat, preconceived system upon it" (Coates 1983: 247). Her findings are

[32] Among these, I should mention also Mindt's (1995) and Kennedy's (2002) contributions, which, however, remain outside the scope of our present purpose.

accompanied by a detailed discussion about the various types of modality which are confined to the two basic semantic categories: root/non-epistemic and epistemic modality.

Depraetere & Verhulst (2008) offer a more promising approach of root necessity. Building on syntactic, semantic and pragmatic distinctions between modal *must* and semi-modal *have to*, the authors propose a better explanation of the 'source' of necessity which they consider as "the driving-force behind the necessity" (Depraetere & Verhulst 2008: 3). More specifically, their aim is to assess the semantic correlation of root modality and source of necessity in the light of the most commonly occurring alternations in present-day English: between subjective necessity with *must* and objective necessity with *have to*, respectively, in the spoken and written sections of ICE GB.

Although the authors (Depraetere & Verhulst 2008: 4) acknowledge that "no particular source is exclusively associated with one particular auxiliary", the notion of 'subjectivity', which is more often associated with *must*, requires a wider perspective. To this end, a binary distinction into 'discourse-internal' (e.g. the speaker in statements, the hearer in questions) and 'discourse-external' sources (e.g. rules and regulations, circumstances, condition etc.) more adequately defines non-epistemic readings. Perhaps the most innovative view of subjectivity refers to the sources which reflect 'personal opinion': e.g. '*in my opinion*', '*I feel*' (Depraetere & Verhulst 2008: 13), and which usually occur with *have to*. Such reinforcing elements remind us of the so-called 'modal harmony' (see Halliday 1970: 331 in Depraetere & Verhulst 2008: 13), and may exert a pragmatic effect on modality as associated with *have to*, in general. Likewise, type of register seems to determine sources, with discourse-internal sources expressed with *have to* being more frequent in spoken data. Therefore, both markers may express a variety of sources ranging from discourse-internal to discourse-external interpretations (Depraetere & Verhulst 2008: 23). Overall, the evidence from this study has potential implications for a similar development of *have to* in New Englishes.

Nokkonen (2006) and Taeymans (2004) are among the few corpus studies that focus on present-day uses of *need* and *need to* as well as *dare* and *dare to*, respectively, the latter pair being of little interest for the present review. While Nokkonen's (2006) inductive examination of the various meanings of *need to* uses two spoken (LLC and

COLT) and two written corpora (LOB and FLOB) of present-day British English, Taeymans (2004) extends the perspective to analyse shifts in the frequency of use in the spoken and written sections of BNC.

To begin with, semi-modal *need to* has been analysed mainly in relation with modal *need* (see Duffley 1994). As a recent innovation in present-day English, Nokkonen (2006: 27) argues for a systematic treatment of semi-modal *need to* similar to the other modal markers such as *must* and *have to*. A slightly different perspective is found in Taeymans' (2004: 97, 99) analysis in which two opposite processes form the scope of her study: the grammaticalization of *need* and the degrammaticalization of *dare*. Her findings suggest that internally motivated *need to* has grown to express through pragmatic inference also external necessity (Taeymans 2004: 108), which alternates with *have to* in negative contexts, but functions as a weakened variant of *must* and *have to* in affirmative contexts. Overall, the author convincingly argues that such a movement from internal to external necessity is consistent with van der Auwera & Plungian's (1998) cross-linguistic findings.

Conversely, Nokkonen (2006) identifies the highest frequency of *need to* in the spoken register, particularly in COLT, and which correlates with first and second person subjects displaying most variation in root readings. Based on Coates' (1983) classification, the data does not exhibit a clear pattern for root readings. However, epistemic meanings with *need to* seem to emerge through conventionalization of implicature (Nokkonen 2006: 63). The study ends with a comparison of *need to* with *must* and *have to* which promises to be a research area of potential diachronic and synchronic variation.

Finally, it seems relevant to add two more PhD-dissertations at the Albert-Ludwig University in Freiburg authored by Friederike Seggewiß (in progress) and David Lorenz (2014), which focus on modality in present-day English. Despite their separate research focus the two studies contribute with further insights on the ongoing grammaticalization of modal items. While Seggewiß focuses on the current changes in the present-day English modality system by analysing spoken corpus-data from a real-time perspective, Lorenz traces the conventionalization of contracted forms such as *gonna*, *gotta* and *wanna* in synchronic data from AmE.

4.2.2 Modal usage in non-standard varieties of British English

Synchronic contrasts of ongoing change and layering within deontic modality in informal dialect data are addressed in Tagliamonte (2004), Tagliamonte & Smith (2006). Trousdale (2003) and Corrigan (2000) offer a wider perspective including both root/deontic and epistemic meanings. These studies confirm that dialect areas represent significant sources in the analysis of change within modality. While Tagliamonte (2004) and Tagliamonte & Smith (2006) approach incipient grammaticalization from a variationist perspective, Trousdale (2003) uses standard sociolinguistic methods.

An analysis of York English in Tagliamonte (2004) and of various other British, Scottish and Northern Irish dialects in Tagliamonte & Smith (2006) reveals that *must* is in decrease across generations and various regional settings. By contrast, *have to* is increasing and takes over the functions expressed by *must* whereas stable variation in the system occurs between *have to* and *have got to*. The multivariate analysis in Tagliamonte (2004) reveals that the reference of the subject as well as type of verb condition variation in the sense of specialization, and which might indicate ongoing change of these semi-modals. Similarly, the results in Tagliamonte & Smith (2006) involve specialization of the two semi-modals according to reference, and which function as a readjustment across dialects.

By contrast, Trousdale (2003) argues that in his corpus of Tyneside English modals tend towards monosemy. Thus, epistemic *must* occurs more often than root readings with this modal, whereas semi-modals *have to, have got to* or *got to* are more often used to express root meanings. Such findings are consistent also in Tagliamonte (2004) and Tagliamonte & Smith (2006). In addition, Trousdale (2003: 281) identifies two tendencies that restructure the modality system in Tyneside English: on the one hand there is simplification, which involves semantic (e.g. increased epistemic *must*) or morphosyntactic restrictions (e.g. root meanings with any variant of *have got to*), whereas on the other hand there is redistribution, which refers to the (re)evaluation of social or linguistic functions of a linguistic form, respectively.

Corrigan (2000) analyses modal use in socio-historical data from a non-standard Irish speech community (South Armagh). Although the choice of the vernacular as well as the research focus (i.e. the non-standard syntagm "modal *be + to*") are different here,

the study stresses that speakers of South Armagh English use *must* as in the contemporary Standard (see Coates 1983) both with deontic and epistemic meaning. Interestingly, *have to* appears to be more frequent in this speech community, whereas *have got to* does not occur at all in the data, but is productively replaced by *be + to*.

Finally, Schulz (2012) investigates the degree of grammaticalization of the domain of past obligation (*had to* and *had got to*) as connected with that of past possession (*had* and *had got*) in British dialect data from FRED. In view of the approach taken, the various paths of change of these expressions are reassessed through Croft's (2000) notion of *intraference* (Schulz 2012: 134). Even though such a focus is not of immediate concern to the present study, it is most relevant in so far as it offers significant insights about the synchronic development of both *have to* and *have got to* in non-standard varieties and, by extension, in other varieties, too.

All in all, the various approaches on modal variation across English dialects seem to converge with respect to one striking aspect, namely that productive expression of deontic/root meaning with *have to* does not occur only in standard English but is quite vigorous in non-standard Englishes, too. Thus, the remarkable use of this semi-modal across several dialects raises questions whether such vernaculars can be defined as truly conservative, after all. However, it would be beyond the scope of our study to elaborate on this issue. Instead, substrate transfer, 'colonial lag' or second-language acquisition represent the most adequate directions when one assess modal usage in dialect areas and other standard educated native and non-native varieties of English (cf. Corrigan 2000: 36-38).

4.2.3 Modal usage in other world Englishes

The discussion on the current trends in modality would be incomplete if we did not include corpus-based approaches focussing on data from other varieties of English.

Leitner (1991), Shastri (1988), Schmied (1994) and Wilson (2005) are a collection of studies on modal use in Indian English. Except for Wilson (2005), who adopts a multivariate quantitative approach on the behaviour of modals across text-types, the former three are largely descriptive contributions focussing on the differences in frequency patterns across written texts. These studies do not specifically concentrate

on obligation and necessity nor have they envisaged addressing the issue of ongoing change, but instead analyse their joint distribution in different text-types of IndE. Even so, valuable information can be inferred on the potential semantic, syntactic and pragmatic tendencies in the use of such expressions in spoken data from ICE-India. A common element shared by all the four authors is the source of data which is drawn from the *Kolhapur Corpus of Indian English* (see section 2.3).

Shastri's (1988: 17-18) study leans on Katikar's (1984) comparison of the core modals in IndE, AmE and BrE, which due to their polysemous nature show no major qualitative differences among these varieties except for the few syntactic confusions in IndE, which arguably reflect "the failure to conform to the sequence of tenses" (Shastri 1988: 17). By contrast, the differences in frequency such as with expressions of 'futurity' and 'hypothesis' indicate "reflections of the peculiar Indian mode of thought" (Shastri 1988: 17).

Leitner (1991) re-assesses subsequently the use of modal verbs in the same database and concludes that the main differences are visible in absolute frequency rates, for example, of *shall* and *should*, and that overall IndE shows a tendency towards increased formality. Although modals in IndE seem to be used as in other native varieties, Leitner (1991: 226) does not exclude potential semantic differences, and calls for more detailed analyses also in the spoken medium. As regards the distribution by genre, specifically in fictional texts, Leitner (1991: 227) suggests that the differences may be related to the general receptivity of such text-types in non-native varieties contrasting with native varieties (see Schmied 1994). Interestingly, information about the use of modal *must* is mentioned in passing in Leitner (1991: 228) referring only to a drop in frequency. Finally, stylistic factors would contribute to the "pragmatic conventions and a particular stage in the development away from the British English ancestor" (Leitner 1991: 229). As regards semantic contrasts between BrE or AmE this study does not offer any further details.

Schmied (1994) analyses several syntactic style features in Indian English in the Kolhapur corpus as compared with LOB, chiefly because BrE is considered still the norm in India. Salient quantitative differences occur with *would* and *might*, which are more frequent in the British data than in the Indian data. Such evidence contradict Shastri (1988), but seem to tally with Leitner (1991). However, specific differences in

obligation and necessity are not part of this study. Nevertheless, Schmied's (1994: 226) assumption that modals are cultural-sensitive is worth taking into account in the present endeavour.[33]

A more sophisticated approach is found in Wilson (2005: 152), who uses multivariate mapping techniques similar to correspondence analyses or factor analyses. The study focuses on intra-varietal variation of modals and their internal consistency across text-types in IndE, BrE and AmE. Since it is predominantly quantitative in methodology, the study does not propose a detailed analysis of modal use, but emphasises rather the way these verbs cluster with specific text-types in each variety. Overall, no clear pattern of obligation and necessity is offered, except that *must* clusters with imaginative writing in the Indian corpus.

In spite of the useful quantitative information, it is surprising that none of these authors thought to analyse semi-modals and their behaviour in relationship to modals. Thus, these studies offer overall a one-dimensional picture on this area of the IndE grammar. All in all, the hypothesis regarding overt cultural differences between IndE and other western varieties can be further explored in the spoken component from ICE-India.

The distribution of obligation and necessity in the currently available ICE-framework has been extensively analysed in Collins (2005; 2008; 2009a; 2009b). His most productive research area is the examination of modal obligation/necessity and modality, in general, in AusE and NZE as compared to BrE and AmE. The uses of modal obligation/necessity in AusE are addressed particularly in Collins (2005). Using data from ICE-Australia which he compares with ICE-New Zealand, ICE-GB and ICE-US, the author observes different rates of frequency in the distribution of root and epistemic readings both in written and spoken registers. While root *must* predominates overwhelmingly in written registers in AusE showing similarities both with BrE and NZE, AmE being at a lower end of this distribution cline, root readings with *have to* and *have got to* in AusE are, as expected, more productive in spoken language endorsing previous observations found in Leech (2003). It is only in the case of *need to* that

[33]According to Tinkham (cf. 1993: 245 quoted in Schmied 1994: 1225), certain Indian features are culture-specific and seem to be a reflection of the speakers' different socio-economic status. Thus, commands are more preferred among non-Westernized speakers of lower socio-economic status than among Westernized Indian speakers with higher status and which are influenced by Anglo-American language use.

Antipodean varieties diverge from the pattern found in BrE and AmE. Drawing on Leech's (2003) hypothesis of 'colloquialisation' of *must*, *have to* and *have got to*, Collins (2005) further assumes 'democratization' as an explanation for the case of *have to* and *need to*. This study suggests that AusE and NZE hold an intermediate position between BrE and AmE.

Despite the valuable regional and register-specific distributional patterns and the functionally-based explanations for the variability in the data, Collins adopts a holistic approach on obligation/necessity. In Collins (2008; 2009a) we notice an increased interest in the register-specific distribution according to semantic reading (e.g. deontic, dynamic and epistemic) of modal items. Again, the analysis provides extensive information of distribution patterns which appear to be consistent across registers. What is felt as missing particularly in Collins (2009a) is a more elaborated analysis of stylistic variability, that has far-reaching implications on theoretical issues, such as semantic change as related to ongoing grammaticalization in apparent-time data from world Englishes.

Collins (2009b) discusses the differential rates of change of modal markers as inferred from their synchronic distribution across ICE representing the inner (BrE, AmE, AusE and NZE) and outer circle varieties (PhilE, SingE, HKE, IndE and KenE). The aim is to compare the distribution and frequency of several modals and quasi-modals in the light of 'colonial lag' (Collins 2009b: 284). Some interesting tendencies emerge particularly in the outer circle varieties. Although quasi-modals are most preferred in inner circle varieties, *have to* is more frequent in outer circle varieties being closer to AmE. Moreover, AusE is closest to AmE, whereas IndE and KenE are described as most conservative in the use of semi-modals. The distribution by register suggests a higher preference of such expressions in the spoken registers among inner circle varieties. The comparison according to register seems to be consistent with the overall distribution patterns where semi-modals are most frequent in AmE, AusE, BrE, NZE and least so in IndE and KenE. Register-specific differences of modal – semi-modal contrasts are more evident between *must* and *have to*, the latter being the most preferred marker in spoken data. The study concludes that AmE is the most advanced variety with regard to the varying distribution of quasi-modals, but also the decline of

modals. Among outer circle varieties IndE exhibits greater stylistic differences between speech and writing.

Biewer (2009) investigates the markers of obligation/necessity in South Pacific second language varieties of English (Fiji English, Samoan English and Cook Islands English) in a corpus of press editorials and press sections from ICE-subcorpora. The study aims to identify similarities and differences of these varieties with AmE, BrE and NZE. They are accounted for in the light of substrate transfer, second language acquisition and exonormative influences (i.e. BrE in Fiji and Cook Islands; AmE but also NZE in Samoa). Notably, such a combined approach moves away from the one-dimensional view of New Englishes as exhibiting 'colonial lag' (cf. Collins 2009b). The frequency analysis displays striking similarities in all the six varieties. Thus *should* is overwhelmingly preferred in all datasets together with *have (got) to*, which is seen as a fixed idiomatic expression, as well as *must* and *need to*. Closer attention is addressed to the uses of *must* and *should*. Similar to other studies, the use of *must* can be related to a decrease found in both BrE and AmE. However, the distribution by semantic contrast reveals that deontic readings are more frequent in FijE, SamE and AmE. Conversely, epistemic meanings are more common in BrE, NZE and CookE. On the other hand, *must* is preferred in passive constructions with a softening effect particularly in FijE and SamE, hence pointing to the complexity of the whole system. Such constructions may be also reflections of second-language strategies and substrate influence as related to social ranking and politeness strategies. An interesting hypothesis is that the patterning of SamE and FijE with AmE in the marked preference for *should* and *must* would reflect further development of the two South Pacific varieties. The connection seems to hold between SamE and AmE but not with FijE. By contrast, CookE patterns with NZE. Despite similar substrate languages the varieties pattern differently as regards their affinity to target models of native speaker.

Mair (2009b) examins the use of modal obligation and necessity in Jamaican English and aims "to assess the synchronic regional orientation of a New English with regard to British or American norms and also to its degree of conservatism" (Mair 2009b: 18). The distribution of such expressions in data from SBC for AmE and conversations from five ICE-corpora (ICE-GB, ICE-India, ICE-Jamaica, ICE-NZ and ICE-Ireland) shows that semi-modal *have to* is most preferred in ICE-Jamaica. The high

preference of this semi-modal is tentatively connected to the common local creole form *hafi* ('have to') possibly functioning also as a constraint for the use of more British *have got to* (see note 15 in Mair 2009b: 28). A similar observation is that *have got to* is also very rare in the Irish data being closer to the AmE pattern found in SBC and which thus confirms the evidence from previous studies (see Corrigan 2000). As regards the use of *need to*, it appears that the Jamaican data patterns well with the American data from SBC. Mair (2009b: 19) comments that such a phenomenon could be linked to the issue of corpus sampling as spoken texts in ICE-Jamaica were recorded approximately ten years later than in the case of the other subcorpora. Among the other corpora, ICE-India is described as highly conservative, particularly in the use of *must*, which matches with BrE, but displays low frequency of the other innovative forms *need to* and *(have) got to*. Overall, Jamaican English exhibits North American profile of variation more than British, and which would suggest a re-orientation towards American speech forms (Mair 2009b: 19).

Building on previous corpus-based research on modality in standard and non-standard English, Tagliamonte & D'Arcy (2007) investigate ongoing structural change in the layering of *must*, *have (got) to*, *got to* and *need to* in a corpus of informal conversations from the Canadian English (CanE) as spoken in Toronto (TorE). The study combines corpus linguistics and variationist sociolinguistics methods to investigate language change (see Tagliamonte 2004; Tagliamonte & Smith 2006).

An analysis in the apparent-time shows that *have to* is the most preferred form particularly among teenage speakers (87%) of TorE. While *need to* is quite stable suggesting specialized use of internally-motivated compulsion, the increase of *have to* is explained in the light of apparent-time evidence (Tagliamonte & D'Arcy 2007: 72, 75). Although *have got to* appears to be in decline, it is striking that *got to* is the second most frequent form in TorE being highly socially marked. By contrast, epistemic *must* still holds position, however it increasingly receives competition from *have to* and *have got to*.

The study assesses the varying frequency rates of deontic *have to*, *got to* and *have got to* by combining grammatical, referential and pragmatic constraints according to age groups. The multivariate analysis reveals that the uses of *have to* and *got to* are consistent across the community and, perhaps most significantly, across age groups.

However, the choice of these forms appears to be conditioned by different social factors, e.g. sex and education. Likewise, *have got to* contrasts in its social evaluation with the other two forms. The main contrast with BrE (see Tagliamonte 2004; Tagliamonte & Smith 2006) is found in the use of *have got to* which is perceived as socially ambivalent among TorE speakers (Tagliamonte & D'Arcy 2007: 82), and which might explain the varying frequency rates between this form and *got to*. The most important finding in this study is the specialization of *have to* and *got to* according to their pragmatic force.

By contrast, Dollinger (2006; 2008) focuses on the development of modal markers in *The Corpus of Early Ontario English* (CONTE) comprising letters, diary entries and newspapers from 1776-1899. The study aims to provide a (post-)colonial perspective on Late Modern English in the light of Trudgill's paradigm of 'new dialect-formation' as well as of linguistic conservatism in CanE. The time frame chosen for the analysis of the variable complex *must* and *have to* is relevant particularly in view of Krug's (2000) claim that the grammaticalization of this semi-modal is dated around the middle of the nineteenth century. The overall frequency analysis of these markers shows that the use of *have to* in CanE is more advanced than in AmE and BrE. Dollinger (2008: 210) comments that such a pattern may be explained methodologically by the lack of AmE data from the nineteenth century. An analysis by periods confirms that the use of *have to* in CanE matches with the AmE pattern, which can be explained as the "import of loyalist speech patterns" (Dollinger 2008: 211). More important seems to be the distribution by genre, where semi-modal *have to* occurs more often in letters from CanE, which might suggest that its spread happened via informal settings. *Must* remains the most productive British variant, while *have to* is steadily increasing during the nineteenth century. Such a development appears to be the result of long-term drift as an independent process in early CanE that cuts across socio-political events of the time frame under analysis (e.g. British-American conflict). As regards semantic contrast, the data displays a bias towards root modality, whereas epistemic necessity is still on the way of its spread. Overall, the real-time analysis shows that CanE is most conservative in the use of epistemic *must*, but is progressive as regards the rapid increase of *have to*.

Last but not least, Noël et al. (2014) provide a first collection of articles that focus on the diachronic dimension of modal usage in four postcolonial varieties of

English, such as AusE (Collins 2014), PhilE (Collins et al. 2014), WAfE (van Rooy & Wasserman 2014) and BSAfE (Wasserman & van Rooy 2014). While AusE and WAfE are established native varieties, PhilE and BSAfE are referred to as non-native varieties. The authors of the four papers seek to establish a historical link between these postcolonial varieties and their parent variety as concerns change in the frequency and semantic profile of various modals and semi-modals. More specifically, the collection emerged from the need to shift the attention of such changes found in BrE and AmE and assess the stages in the development of postcolonial varieties (e.g. Schneider 2007) against the backdrop of diachronic data from corpora on New Englishes (Noël et al. 2014: 3-4). The data used for these studies covers the eighteenth century to the present (AusE, WSAfE), the 1950's and early 1960's which is compatible with the Brown corpora (Phil Brown, BSAfE) and the 1990's as covered in ICE (PhilE, WSAfE).

In Collins (2014), data from AusE reports the general tendencies as regards the frequency of expressions marking obligation/necessity and volition/prediction. The differences are found mainly in the individual use of items. Thus, modal *must* has sharply declined its frequency in AusE, which suggests an independent evolution from the two reference varieties. By contrast, semi-modals show to be lesser used in AusE.

Collins et al. (2014) study the recent changes in the use of six modals and semi-modals in the Phil as compared to BrE and AmE. The analysis shows that PhilE uses a divergent pattern as compared to BrE and AmE, which the authors refer to a phase of "endonormative stabilisation". Overall, the decline of modals in PhilE is sharper (even though these continue to be preferred in press reportage), whereas semi-modals increase at a higher rate than in AmE.

Wasserman & van Rooy (2014) evaluate diachronic change in modality in WSAfE as compared to other native varieties (BrE, AmE and AusE) as well as the influence that contact with Afrakaans has had on its development. It appears that modal *must* is more frequent in WSAfE whereas *have to* shows a lower rate of increase in use, even though it follows the general trend as other varieties towards colloquialization. What stands out as remarkable in WSAfE is the near equivalence in the frequency of *must* and *should* both in written and in spoken language, which is also indicative of their polysemy. The authors stress the conservative profile of WSAfE, which is reported to have developed differently from the other native varieties. In addition, Wasserman &

van Rooy (2014: 41) consider the influence of Afrikaans as essential in the divergent distribution of modal *must*.

Finally, van Rooy & Wasserman (2014) compare diachronic data of BSAfE (a non-native variety) with WSAfE (essentially a native variety) to find out if the two varieties converge in the frequency of modals and semi-modals. The most general finding is that, while WSAfE reports a decline in the modals and semi-modals (see Wasserman & van Rooy), BSAfE shows little change in the second half of the century. Another difference between these varieties is reported in the case of the selection between *must* and *should*, which in WSAfE are overlapping in meaning whereas in BSAfE the modal *must* is, overall, more frequent. Such a state of art suggests a different dynamic of change in BSAfE as compared to WSAfE, but also to other non-native varieties of English (e.g. PhilE).

4.3 Data and methods of analysis

The present section introduces the data sources and methods of the corpus analysis. Methodological considerations will focus on two aspects: (a) the register- and genre-specific text categories in ICE and (b) the descriptive approach using the comparative corpus linguistics method in the apparent-time framework (Leech et al. 2009).

4.3.1 The corpora

The main data source for this study consists in spoken material selected from a collection of each 180 private and public dialogue samples of adult educated speakers of English of the recently compiled components of the *International Corpus of English*

(ICE).[34] The corpora of main interest are ICE-JA,[35] ICE-India and ICE-Ireland.[36] The British component ICE-GB will serve as reference for the parent-variety.

The underlying principle of the framework of ICE is to facilitate comparative studies on among world Englishes. The main asset of such a database is its uniform design of balanced structures of text categories which represent both the spoken and the written medium. The common design of the 1-million words corpora comprise 500 texts (300 of spoken and 200 of written medium) each of approx. 2,000-words per text (see Nelson 1996a). Each component uses a common markup system both for spoken and written texts (Nelson 1996b; Nelson et al. 2002: 9-13). In addition, the British component ICE-GB is delivered with its own retrieval software ICE CUP (see Nelson et al. 2002), and is therefore the only fully tagged and parsed subcorpus.[37]

Within spoken texts a distinction is made between dialogues and monologues. Several parameters were considered in the sampling process: the private or public context of communication (i.e. presence or absence of audience), direct vs. distanced context (i.e. face-to-face conversation or use of technical means (telephone, tape-recorder etc.)), level of formality (formal vs. informal/casual), spontaneous vs. prepared for the case of spoken public monologues (see Greenbaum 1991; Schmied 1990). Further distinctions refer to written texts: scripted (written to be spoken), non-printed (handwritten or typed) and printed. Although the ICE project aims at conformity, it has proved impossible to match speakers and writers exactly in all the parallel corpora. For

[34] Initially, the project design of ICE aimed to include contributions of spoken material and printed texts from eighteen countries where English is the first or an official additional language (Greenbaum 1991; 1992; 1996). These are: Australia, Cameroon, Canada, Caribbean (Jamaica), East Africa (Kenya, Malawi, Tanzania), Fiji, Ghana, Great Britain, Ireland (Eire and Northern Ireland), Hong Kong, India, New Zealand, Nigeria, Philippines, Sierra Leone, Singapore, South Africa and the USA. To these, further research teams have joined the project such as Malta, Pakistan, Trinidad and Tobago or Sri Lanka. See webpage http://ice-corpora.net/ice/.

[35] The Jamaican component of the International Corpus of English (ICE-JA) was compiled at the Albert-Ludwig University of Freiburg in cooperation with the University of the West Indies (Mona/Jamaica). For details on the recording conventions for some of the conversations (text files S1A001-S1A040), see Deuber (2009: 6-8). A further observation is necessary: At the initial stages of undertaking the present study an earlier version of the Jamaican corpus dating from May 2007 was used, which is different in certain ways from the official version. Yet, even though for reasons of incompatibility with the present conventions some of the dialogues from the earlier corpus were replaced in the official version, the frequency of the four expressions is minimally affected by these changes.

[36] The Irish component of ICE is sampled to represent equally both the variety spoken in Northern Ireland as well as in the Republic of Ireland (Kallen & Kirk 2008: 4).

[37] ICE CUP stands for the *corpus utility program*, which is an advanced tool to explore the corpus. For the present study ICE CUP version 3.1 was used.

this reason biographic information and subject-matter of some text categories could not be fully matched.[38]

Additionally, a narrow set of sociolinguistic variables, such as biographic data of the speakers (age, occupation/educational background, sex), channel/means of conversation (direct, distanced) including bibliographic information about each text categories, allows the studying of socio-stylistic variation between native and non-native varieties. However, some of the variables, i.e. age and sex, appear to be problematic if we consider including them as measuring effects for the use of modal obligation/necessity. To begin with, speakers are differentiated according to age-groups ranging 18-25, 26-45 and >46. Furthermore, spoken texts contain recordings of conversational settings among participants which are either males or females or which occur as a mixed set of both sexes. Since only ICE-CUP tool provides automatic retrieval of biographic information, a manual extraction in ICE-JA, ICE-India and ICE-Ireland has proved to be rather time-consuming at a more advanced stage of this study. Arguably, such limitation may not sustain a systematic study in apparent-time. However, considering that ICE has been virtually compiled within the same time frame and contains texts from more than one speaker from different regional settings ultimately warrants such an approach (Mair 2009a: 1118). Therefore, in this study regional setting and level of formality (formal vs. informal) count as sociolinguistic variables.

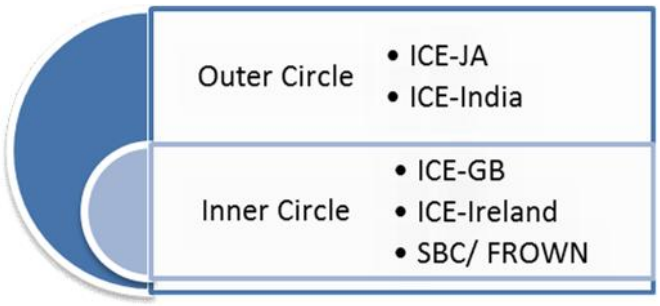

Figure 4.1. Comparison between Inner and Outer circle varieties

[38] These and other problems in the compilation process of ICE are discussed in Holmes (1996), Leitner (1992), Mair (1992) and Schmied (1990; 1996). Furthermore, one should not overlook some inherent differences between the three varieties: e.g. Jamaican English with basilect and mesolect speech forms that permeate the educated standard (i.e. acrolect) variety more often occurring in 'conversations' (Deuber 2009: 6). According to Deuber (cf. 2010: 109), such features reflecting the creole-continuum are characteristic of the two existing Caribbean ICE corpora so far, ICE-JA and ICE-T&T.

Analysing the four modal markers only against the British component, ICE-GB, yields a one-sided picture especially if one aims to find out whether JamE, IndE or IrE still follow British norms of usages rather than North American. Although an equivalent ICE-US component is still lacking, for comparison I also analyse the *Santa Barbara Corpus of Spoken American English* (SBC) with approximately 249,000 words of spoken texts. Analogously to the configuration of ICE, SBC contains recorded face-to-face conversations with speakers from different regional settings of the United States, of different age, occupations, or ethnic background, providing thus a common ground of the spoken register from the other corpora. Moreover, it is intended to fill the spoken part of the missing ICE-US component and thus can be considered as a suitable approximation. However, for various reasons SBC is not a perfect match: first of all, it differs from ICE in terms of the number of dialogues and in the number of words per text. Also, the dialogues are not classified according to the same design as found in ICE. Nevertheless, it remains a useful tool to assess variation in terms of relative frequency in the joint distribution of these items. Likewise, I will use data from Frown for American news reports (see next section 4.2.2).

To sum up, as Figure 4.1 shows, the present study focuses on patterns of use with strong obligation/necessity in the outer circle varieties (JamE, IndE) which will be compared with the inner circle (BrE, AmE and IrE).

4.3.2 Selected text categories in ICE

The spoken section of ICE contains eight text types which are divided into private and public dialogues. The total size of the present database for each subcorpus is of approximately 360,000 words. Table 4.3 lists the text subcategories which are used in the analysis: ‘face-to-face conversations’, ‘phone calls’, ‘class lessons’, ‘broadcast discussions’, ‘broadcast interviews’, ‘parliamentary debates’, ‘cross-examinations’ and ‘business transactions’. The number of texts for each category is specified in brackets.

Dialogues (180)	**Text category (S1)**	**Text code (S1A-S1B)**
Private (100)	*Conversations* (90)	S1A-001 *to* S1A-090
	Phone calls (10)	S1A-091 *to* S1A-100
Public (80)	*Class Lessons* (20)	S1B-001 *to* S1B-020
	Broadcast Discussions (20)	S1B-021 *to* S1B-040
	Broadcast Interviews (10)	S1B-041 *to* S1B-050
	Parliamentary Debates (10)	S1B-051 *to* S1B-060
	Cross-examination (10)	S1B-061 *to* S1B-070
	Business Transactions (10)	S1B-071 *to* S1B-080

Table 4.1. The spoken text categories in ICE. Source: http://www.ucl.ac.uk/english-usage/ice/design.html.

Although the spoken conversations represent the primary data source, the assessment of a small number of written texts allow for a more accurate description of the extent of either innovation as indigenisation or of 'colonial lag' in the present data. It is often in post-colonial settings where the norms of the written mode are considered as the target of 'educated' English and which vary considerably from the local variety (e.g. India, Jamaica) (see 4.5.3). Therefore, the research design includes in addition to spoken dialogues also 20 press reports from each ICE subcorpora (see Table 4.4):

Register	**Corpus**	**Text categories**	**Text code**
Spoken	ICE *(360,000 words)*	Private dialogues	S1A-001 *to* S1A-100
		Public dialogues	S1B-001 *to* S1B-080
	SBC *(249,000 words)*	Private dialogues	SBC001 *to* SBC014 (*Part 1*)
			SBC015 *to* SBC030 (*Part 2*)
			SBC031 *to* SBC046 (*Part 3*)
			SBC047 *to* SBC060 (*Part 4*)
Written	ICE *(40,000 words)*	Press News Reports	W2C-001 *to* W2C-020
	Frown *(approx. 88,000 words)*	Press reports	A01 *to* A44

Table 4.2. Register selection for the present study

Due to the similar time frame (1991) of compilation with ICE and to substitute the missing written AmE component, a total of 44 press texts of approx. 88,000 words of text are taken from the Frown corpus.

Most often, the term 'register' is associated to non-linguistic/situational aspects.[39] For example, Biber et al. (1999: 15) use the term 'register' in a broad sense as the "situational characteristics such as mode, interactiveness, domain, communicative purpose, and topic […]", whereas Biber & Conrad (2009: 6) narrow it to "a variety associated with a particular situation of use (including particular communicative purposes)". As for the present case, grammatical variation across New Englishes will be assessed as intertwined with register differences. According to Biber et al. (1999: 15-17), variation at grammatical level occurs at the intersection between four major registers which cover a wide range of situational contexts: 'conversation', 'fiction', 'news', and 'academic prose'.[40] For example, in face-to-face conversation interlocutors engage in interaction by sharing a personal communicative purpose. By contrast, news reports are written for a wide range of readership and do not involve direct interaction. In view of the chosen data, both conversation and news seem to cover the same regional and social characteristics (Biber et al. 1999: 16), so that the two registers represent two ends on the formality cline, with 'conversation' situated at the lowest informal and 'news' at the highest and most formal end. More specifically, informal style which reflects spontaneous speech habits contrasts with the formal style which is more elaborate.[41]

Although the formal end appears to be fairly homogenous, further distinctions can be made. For instance, it is likely that 'class lessons', 'broadcast discussions' as well as 'parliamentary debates' share to some extent the same dialogue style even though the former are usually less interactive and represent rather the semiformal segment of conversational settings (Deuber 2010: 109, 113). By contrast, interaction in 'cross-examinations' and 'business transactions' involves highly specialized discourse,

[39] For a detailed survey of previous definitions of 'register' as well as the related terms 'genre', 'text type' and 'style', see Biber (1994: 51-53). For most recent explorations on syntactic variation in different genres, see the edited volume by Dorgeloh & Wanner (2010).

[40] This can be explained by the fact that linguistic choices are linked to various circumstances of speech production generating a wide range of variation patterns (Biber et al. 1999: 20-21).

[41] In view of Biber's (1988) feature analysis, the selected registers cut across Dimension 1 of 'involved' vs. 'informational' language production and Dimension 3 of 'elaborated' vs. 'situation-dependent' reference. See also Koch & Oesterreicher's (1985) conceptualization of the written vs. spoken duality in language as 'language of proximity' (*Sprache der Nähe*) and 'language of distance' (*Sprache der Distanz*). According to their elaborate distinction, the two poles *spoken* and *written* form a continuum which represent various realizations of the linguistic expression (Koch & Oesterreicher 1985: 18). In a similar vein, on the basis of Lyons's notion of *medium-transferability* of language, Esser (2000) argues for a general framework of presentation structures in speech and writing, i.e. stylistic choices encoding at the levels of form and substance.

which is a characteristic of professional settings (Giltrow 2010: 29; Schilling-Estes 2002: 375).

To sum up, the present study focuses on a narrow stylistic range, which consists of informal (i.e. low level) and formal (i.e. high to medium level) conversation settings in addition to written press reports. The validity of the assumption that the asymmetries in the use of modal obligation/necessity in ICE indicate stylistic preferences remains to be clarified in the following chapters.

4.3.3. Data extraction: framing the variable context

The extraction of the relevant tokens from the bundle of data was achieved with the AntConc KWIC concordance list[42] by generating several queries with all the morphological variants of *must*, *have to*, *(have) got to* and *need to* including contracted forms (‘*ve to*, ‘*ve got to*, ‘*s got to*). To ensure variability, for the more detailed analysis each token was restricted to present tense, affirmative, declarative utterances and examined in the light of the NICE properties (see 4.6.2). Relevant to the analysis is that the counts for *have to* and *have got to* include some particular uses, such as adverb interpolation (e.g. *have* still *to*, *have* only *to*), and which are commented also in Smith (2003: 248). By contrast, cases in which perfective uses of *have to* or *(have) got to* are part of expressions (e.g. *have to/(have) got to do with sth./sb.* or *have to say*) as in (11) were not considered in the analysis:

(11) So I have <,> I have always benefited from what Chomsky **had to** *say* (ICE-India.S1A-082.txt)

According to Smith (2003: 248), such uses of *have to* are not linked to necessity but represent a different direction of development of possessive *have + to*, which in the present dataset can be considered as an instance of ongoing change (see Krug 2000).

[42] Apart from ICE CUP which is incorporated with ICE-GB, I used the concordance programme AntConc version 3.2.1 which can generate concordance lines, concordance plots, or provides a set of tools to analyse clusters, word frequencies and keywords. AntConc is a freeware concordance programme which has been developed by Laurence Anthony at the Waseda University, Japan. See http://www.antlab.sci.waseda.ac.jp/software.html.

Additionally, it should be added that such ambiguity arises only in relatives, interrogatives or other cases of wh-fronting.

The following invariant contexts were excluded from the analysis of semantic contrast (see Chapter 6):

- Any past tense forms of *have to* or *need to* in (12);

(12) So I didn't have one of the typical stories where after school I couldn't get a job where it was hard and I had to be job-hunting and sending resumes here and there <#>I *didn't* **have to** go through that (ICE JA. S1A-058.txt)

- negated forms, as in (36) and (37), as there is no semantic equivalence between negated *must, have to, (have) got to* nor *need to*:
- Interrogatives as in (13) and (14):

(13) But what what **must** be the criterion for this? (ICE-India.S1a-013.txt)

(14) Why you **have to** go down to the Union? Because that's where the money is (ICE-JA. S1A-034.txt)

- Non-finite forms (15) including future forms (16) or hypothetic constructions (17);

(15) They come number one from those who have never had to set foot through the gates of this campus and therefore do not know what it is like **to have to** be studying and **to have to** be worrying about how you're gonna be paying your fees (ICE-JA.S1A-048.txt)

(16) I mean there's a road but if you go on the road you'**ll have to** go like twenty miles cos you'**ll have to** pass the area and come back to reach it by the main road (ICE-JA.S1A-024.txt)

(17) See I'**d have to** suggest to you that the length of time that's likely for you to have completed that manoeuvre would be something in the region of twelve seconds (ICE-Ireland.S1B-061.txt)

- formulaic expressions as in (18) and (19) because of their inherent performative function, also called 'discourse rituals' (Tagliamonte 2007: 73);[43]

[43] Despite their conventionalized use in spontaneous interaction these structures prove to be a good source when searching for variation. What seems to be striking is that the two outer circle varieties continue to exhibit a preference towards using *must* (ICE-India total=24 and ICE-JA total=25) rather than semi-modals, thus preserving in this way the 'conservative' colonial linguistic heritage. Conversely, *have to* in formulaic expressions is overall more frequent in the British and Irish data whereas more recent innovations such as formulaic *got to* appears only once in the Irish conversations. In addition, there are four occurrences of *need to* in the Jamaican corpus.

(18) I **must** say that uh professionally I have been English Studies Officer of the British Council (ICE-India.S1b-071.txt)

(19) I **have to** share a sentiment expressed by one of your members that leave is something that should be encouraged (ICE-JA.S1B-079.txt)

- Further exclusions refer to unfinished or elliptic utterances:

(20) <S1A-020$B> <#> He he's so what was his craic something like you **have to** or <S1A-020$A> <#> Well he said you know och well (ICE-Ireland.S1A-020.txt)

(21) <ICE-IND:S1A-038#119:1:B> From there you **must** I mean you always <ICE-IND:S1A-038#120:1:A> I have work (ICE-India.S1A- 038.txt)

- Finally, hesitations and repetitions as well as unclear strings of conversation were excluded following the markup system in ICE (Nelson 1996b: 41).[44]

Even though this filtering process considerably reduces the number of instances to be analysed, it is a necessary process for the description of variation patterns. Thus the focus is on semantically motivated (or 'nonsyntactic') instances – as opposed to 'syntactic' ones (Myhill 1995: 166)[45] – in which such variation is possible at least in principle.

In order to avoid premature conclusions, the interpretation of quantitative measurements is counterbalanced with qualitative evaluations of the corpus data (Leech et al. 2009: 32; McEnery & Wilson 2003 [1996]: 76-77). Wherever useful, significance will be assessed on the basis of descriptive statistical tests, such as the two-way chi-square test (Meyer 2002: 130; Oakes 1998: 24-26). In cases of low frequency counts such analyses will not be performed. In order to avoid skewed statistics and thus

[44] Such exclusions refer to the following annotations: <?>…</?> for uncertain transcriptions; <.>…</.> for incomplete words; <X>…</X> for extra-corpus text; <foreign>…</foreign> for foreign words; <indig>…</indig> for indigenous words; <unclear>…</unclear> for unclear words (Nelson 1996b).

[45] Following Myhill (1995: 166), equivalence between *must* and *have to* with past reference is syntactically motivated in constructions with perfect aspect. These uses are always epistemic as in (1) and (2):

(1) She **must** *have called* at least three times you know after that (ICE-India.S1A-091.txt)
(2) <ICE-IND:S1A-028#240:1:B> As as you have American started earlier English <indig> na </indig> I think we should although we <,>I think we already have a Indian style we have it you know ICE-IND:S1A-028#241:1:A>We **have to** *have considered* a limit na extreme limit (ICE-India.S1A-028.txt)

The majority of examples of perfective aspect with epistemic meaning are with *must*, while the construction with *have to* in the present perfect as in (2) is rather an exception and appears only once in the IndE as well as in BrE data.

enhance the value of predictability, significance will be reported only if the value obtained for the level of probability $p \leq .05$.

4.4 Structural and sociolinguistic dimensions of changes in modal usage in World Englishes

Three possible explanations have been suggested for the varying degrees of changes in the domain of strong obligation and necessity particularly in BrE and AmE. The structural change commonly assumed to be at work is grammaticalization (see 4.3.1), which links morphosyntactic change to semantic developments (Hopper & Traugott 2003; Krug 2000; Traugott 1989) or to the shifts in discourse frequency as part of social diffusion (Tagliamonte 2004; Tagliamonte & D'Arcy 2007; Tagliamonte & Smith 2006).

Apart from changes at structural level, the recent trends in modality seem to be related to innovations in discourse as well as to socio-cultural processes (see 4.3.2). One explanation is the 'Americanization' of English as suggested in Mair (1997; 1998) and Leech (2003). Another possible explanation is the so-called 'colloquialization' of English (Mair 1997; 1998) which is related to the former (see Leech 2003; Mair & Leech 2006). Associated to this trend, the 'democratization of discourse' is used in Myhill (1995) and Smith (2003) to refer to less marked inequalities of power as reflected in public discourse.

Of further relevance to the study seem to be substrate influence and prescriptivism (see section 4.3.3). For the case of JamE, the local creole variety has a potential role in determining variation. IrE, historically a language-shift variety displaying Gaelic influence to various degrees, and Indian English an 'associate language' on the multilingual Indian sub-continent are also special cases of language-contact. As regards prescriptivism, it is precisely in multilingual communities (i.e. the Indian subcontinent) that correct language use is consciously prescribed to promote standard speech conventions (cf. Leech et al. 2009: 263).

The following sub-sections (sections 4.4.1, 4.4.2 and 4.4.3) discuss these factors in some more detail as potential determinants for ongoing change in spoken data from ICE-JA, ICE-India and ICE-Ireland.

4.4.1 Grammaticalization: explaining ongoing semantic and structural change

Although modals and semi-modals in present-day English are already grammaticalized expressions their meaning and use has extended to various contexts in discourse. Regardless of the research perspectives adopted on this topic, modal auxiliaries are considered as a paradigm case of grammatical change.

Given such variability in use, other questions for the present study are: How can the behaviour (syntactic, semantic and discourse-pragmatic) of obligation and necessity be accurately described in ex-colonial Englishes as ongoing change? And, are the mechanisms of grammaticalization relevant determinants in identifying ongoing semantic changes of the semi-modals *have to*, *(have) got to* and *need to* as compared to modal *must*? Although grammaticalization is confined mainly to historical processes of grammatical change, variation patterns can be analysed synchronically as representing one specific stage of ongoing development (Hopper & Traugott 2003: 2).

As a research framework, grammaticalization links the diachronic and synchronic dimensions of language change as continuous in the sense of "fluid patterns of language use across time or at a synchronically segmented moment in time" (Traugott & Heine 1991: 1). More recent utterance-based accounts on language change (Croft 2000: 63) approach grammaticalization as a complex mechanism which incorporates both ongoing and historical changes as panchronic in the sense of Heine et al. (1991: 261)[46]. It is such a functional perspective which is taken as point of departure for the analysis of semantic intricacies within modal expressions in the New Englishes under scrutiny.

4.4.1.1 Unidirectionality, semantic change and subjectification

In the traditional view, grammaticalization is the process in which an open-class element develops into a closed-class element involving several formal and semantic changes, such as loss of autonomy in exchange of syntactic bondedness through reanalysis, phonetic reduction or loss in meaning (Fischer & Rosenbach 2000: 2).

[46] Ideally, a panchronic view of grammaticalization comprises diachronic change and synchronic variation. However, according to Heine & Narrog (2010: 22), "there is so far no general theory of language that appropriately accounts for panchrony."

Typically, such changes from less to more grammatical functions are assumed to be predictable evolving in conceptual **chains** (Heine 1993: 58; Heine et al. 1991: 220-228) which illustrate internal relational patterns that channel the change; or on pre-defined **clines** (Hopper & Traugott 2003) of structural transitions. A fundamentally acknowledged principle governing the whole framework is the *unidirectionality* of such developments.[47]

The several clines which have been proposed resemble implicational hierarchies with semantic change as part of the unidirectional progression from more semantic to less semantic or pragmatic meaning (Traugott & Heine 1991: 4). A common characteristic of such clines is increase in abstractness on which a linguistic item becomes more grammatical as in (22):

(22) content item > grammatical word > clitic > inflectional affix (Hopper & Traugott 2003: 7).

Such a cline describes primarily structural changes in an upwards movement with the items represented at the left being least grammatical than those at the right. For the evolution of modal auxiliaries the following unidirectional clines involving additional changes in meaning have been suggested:

(23) Facultative (or dynamic) > deontic > epistemic > futurity, conditional etc. (Goossens 1987: 118)

(24) MV (Main verb) > PM (premodal) > deontic > weak epistemic > strong epistemic (Traugott 1989: 43)

[47] Recent research has challenged the concept of grammaticalization as a self-contained theory, as well as the principle of unidirectionality as an essential component in understanding the various stages involved in processes of language change. For extensive critical assessments of the various positions, see Campbell and Janda (2001), Fischer et al. (2004), Lass (2000) and Newmeyer (1998). By contrast, Haspelmath (1999: 1054-1055) proposes a usage-based explanation that defends the irreversible and unidirectional character of grammaticalization as motivated by a number of "ecological" factors: unconscious processing, routinization, maxims of action and the invisible hand processes. In addition, Fischer (2003) critically evaluates the principle of unidirectionality, the idea of conceptual chains and parameters of grammaticalization as determinants of language change both in formal and functional theories. The debate, in fact, originates in the fundamental differences as to how such phenomena are perceived within the two linguistic orientations. While generative scholars disregard the aspect of variation and of diachronic development in terms of conceptual chains, functional linguists approach this issue from a different angle as they set out to capture the features which determine grammatical changes (Fischer 2003: 446-451).

Both clines reconstruct the historical stages and mechanisms involved in semantic change from pre-modal to modal meanings as desemanticization (Goossens 1987: 119) in (23) or as the extension from deontic/root to epistemic modality through *subjectification* in (24). From synchronic perspective the effects of such unidirectional changes are identified in the layering of older (i.e. modal auxiliaries) and newer variants (i.e. semi-modals) with variation functioning as the outcome of such variability.

Cross-linguistic evidence, in addition, suggests that certain semantic properties of modality are shared among other languages of the world and evolve along specific **paths** of change from pre-modal to modal (root/agent-oriented and epistemic) and post-modal meanings. A common feature of such pathways is generalization of meaning as metaphorical extension of 'grams' (Bybee et al. 1994; Bybee & Pagliuca 1985), which suggests that modal agent-oriented meanings (terminology taken from Bybee et al. 1994) emerge earlier than epistemic ones. Other authors (see van der Auwera & Plungian 1998) have developed semantic maps which trace developments of meanings from cross-linguistic typological perspective by emphasising the non-epistemic (i.e. 'participant-internal', 'participant-external') or epistemic sources of modality. According to such representations, the shift from internal to external necessity is consonant with the unidirectionality claim.

Several principles and parameters underlie the process of grammaticalization, such as Lehmann's (1995 [1982]: 123) parameters, which are classified according to paradigmatic and syntagmatic aspects: i.e. the combination and selection of a linguistic sign. In addition, Hopper (1991) identifies a set of five "principles of grammaticization": *layering*, *divergence*, *specialization*, *persistence* and *de-categorialization*. Heine (1993: 54-58; 2003: 579; 2010: 405) proposes four parameters operating at all dimensions of language organisation. A recurrent mechanism claimed to define grammaticalization is 'semantic bleaching' which he refers to as *desematicization*. The effect of this process is loss (or generalization) in meaning content. Further parameters are *extension*, which is identifiable in the rise of new grammatical meanings in new contexts (context-induced reinterpretation); *decategorialization*, which leads to loss in morphosyntactic properties characteristic of lexical or other less grammaticalized forms; and *erosion* ("phonetic reduction") as loss in phonetic substance.

More recently, emphasis has been placed on the emergence of new functions of linguistic items out of pragmatic inference in context (see Heine 2003; Hopper & Traugott 2003). In the same vein, Traugott & Dasher (2002) claim that semantic change is triggered by pragmatic-semantic factors.[48]

It has been suggested that epistemic meanings are *metaphorical* mappings of real-world representations on the internal intellectual and psychological states (Sweetser 1990) or that they involve conventionalization of implicatures as *metonymic* extensions (Traugott & Dasher 2002). Another mechanism is *subjectification*, which from the perspective of grammatical change refers to the development of new meanings of lexical material involving the speaker's/writer's judgement or perspective on the proposition (see Traugott 1988; 2003; 1997; 1989; 1995; 2010; and Traugott & Dasher 2002).[49] Thus, the development of epistemic modals is a case of diachronic subjectification (see Traugott 1989) that strengthens the unidirectional hypothesis.

A detailed approach on the subjectification of modal *must* is found in Goossens (2000), who claims that modal shift, i.e. meaning extension, from 'participant-internal' to 'participant-external' to epistemic modality has followed several minimal steps which cannot be explained neither through metaphor nor through metonymy. Instead he suggests that such shifts occurred through "parallel chaining" of modal uses moving in the same direction, and which is inspired from Langacker's (1987: 69 cited in Goossens 2000: 150) cognitive notion of "partial sanction". Goossens concludes that the shift to epistemic meanings was initiated within deontic modality in the interaction between parallel chaining and increased subjectification.

Perhaps most relevant for the present purpose is the treatment of the opposition subjectified – non-subjectified use of *must* and *have to* in present-day English (Goossens 2000: 163). Assuming the existence of prototypical meanings of the modal *must* (participant-external, deontic and subjectified epistemic meanings), the corpus data provides evidence of so-called transitional ("inferential") uses of this marker, which are

[48] Other mechanisms are *reanalysis* and *analogy*, though these are not restricted solely to grammaticalization. These two types of changes are associated particularly with generative oriented linguistics (see Fischer 2003).

[49] Unless subjectification arises out of semantic change as the result of formal reanalysis, which emerges into new grammatical functions, this specific mechanism need not be necessarily related to grammaticalization (Cuyckens et al. 2010: 6). By contrast, Langacker (1990; 1998; 1999) refers to subjectification as meaning change, e.g. in the English modal auxiliaries, as similar to the force dynamic-model adopted in Talmy (1988) and Sweetser (1990) (see review in Narrog 2010: 387-389).

not yet subjectified. Apart from epistemic uses, subjectified meanings refer also to the speaker's authority and, more important, to "rules and instructions to which the speaker gives his/her support, or which he submits to the hearer on the basis of his expert knowledge" (Goossens 2000: 164). The conclusion of this study is that subjectification within deontic modality has initiated the shift to epistemic uses through successive chaining, which, ultimately, favoured subsequent conventionalization of certain forms (Traugott & Dasher 2002: 125-129).

According to Traugott (1995: 46), subjectification deserves increased attention in the study of semantic change as it appears to be a "characteristic of all domains of grammaticalization". Likewise, Traugott & Dasher (2002: 30) state that it is "the most pervasive type of semantic change to date". Such an approach allows for a more accurate description of both diachronic evolution and also synchronic changes within modal necessity.

4.4.1.2 The role of frequency

A typical problem for the corpus linguist consists in identifying phenomena of grammaticalization with any instance of language change in general. A widely acknowledged view about processes of grammaticalization is the existing relationship between grammatical status and increase of discourse-frequency (see Bybee et al. 1994; Hopper & Traugott 2003). As shown in previous corpus-based studies on modality in BrE and AmE, a first indicator of ongoing change can be found in the diverging frequency patterns across registers. Hence, some of the most challenging methodological concepts in corpus linguistics refer to discourse frequency, low-frequency items, representativeness of corpus data including the register and stylistic dimension, and their statistical significance (see Hoffmann 2004; Mair 2004).

According to the utterance-based functional perspective on language change, one way to identify grammaticalized forms is to highlight frequency of utterance in terms of "the replication of linguistic structures in utterances in language use" (Croft 2000: 7). Seen from this perspective, the outcome of replication is closely linked with innovation through alteration of linguistic rules and structures spreading gradually across the linguistic system (Haiman 1998). Bybee's (2003) definition of grammatical

change as based on frequency relates language form to function, with grammar arising out of the conventionalization of utterance through repetition:

> […] I will argue for a new definition of grammaticalization, one which recognizes the crucial role of repetition in grammaticization and characterizes it as the process by which a frequently-used sequence of words or morphemes becomes automated as a single processing unit. (Bybee 2003: 603)

Additionally, Bybee (2003: 604-605) distinguishes between type and token frequency and argues that grammaticalization is initiated with the increase of type frequency. Unless token frequency increases the generalization of meaning, the new grammatical form cannot spread in the system. However, Mair (2004) draws attention on the caveats of such a line of argumentation as increase in discourse frequency is not always related to grammaticalization. The same holds for the whole process itself, as not all (semantic) changes are proof of emergent grammatical structures (see Traugott & Heine 1991).[50]

For the present purpose frequency of use has to be correlated with a particular type of meaning (i.e. root or epistemic) so as to identify form-function instability. Thus, the methodological point of departure for the analysis will be the identification of semantic-pragmatic aspects as determinants of change.[51]

4.4.1.3 Diachronic evolution of must, have to, (have) got to *and* need to

Ongoing change of strong obligation and necessity in JamE, IndE and IrE must be seen against the historical evolution of the four verbal constructions. Historically the oldest among the four expressions, the modal *must* was originally a past tense marker *motan* which started to exhibit auxiliary properties as early as Old English (Traugott 1989). According to Warner (1993), starting with the Middle English period it is used as an epistemic marker. Later on, from the late Early Modern English period onwards the emergence of the semi-auxiliaries *have to* and *have got to* complete the system bearing the role of functional equivalence. Although associated with negative connotation, first occurrences of *(have) got to* into English go back to the beginning of the nineteenth century (Krug 1998a: 178; 2000: 63). In regards to the grammaticalization of semi-

[50] See also Kuteva (2001: 167f) with regard to the difficulty to identify the mechanisms which lead to semantic change or 'context-induced reinterpretation' in grammaticalization.
[51] Such a point of view concords with Traugott and Dasher's (2002: 87) claim that "[g]rammaticalization, like semanticization, involves interaction of linguistic structure and language use."

modal *have to* syntactic re-analysis has been accounted to have explanatory value (Brinton 1991; Krug 2000). This does not seem to apply for *(have) got to*, which appears to be rather the result of frequency effects and of processing the constraints of a multi-constituent structure (Krug 2000: 63, 183). Apart from high discourse frequency, which is an essential factor in the emergence of the obligation meaning with *(have) got to*, Krug (2000: 64) does not exclude the possibility of semantic analogy with the older form *have to*, which might have facilitated the coalescence of adjacent linguistic items such as pronouns (e.g. *I've got to*). Considering that the discourse frequency of *have got to* is generally relatively low across the datasets under study (see Chapter 5) – especially when compared to *have to* – analogy to *have to* appears to be the more important factor in the spread of this marker in the New Englishes. However, in contrast to the variant form *have to*, coalescence with *have got to* involves further processing constraints, such as interpolation of *got*, which in Krug's (2000: 64) opinion function as conflicting forces in language change. Also relevant in this context is Plank's (1984) claim that *to*-contraction is indicative of auxiliary status for *have to*. Similarly, Fischer & Rosenbach (2000: 18) argue that the grammaticalization of possessive *have* into an auxiliary was caused by change in word order from SOV (OE) → SVO (ME) with infinitive *to* as an adjacent in all contexts. Fischer (2003: 460), in addition, argues that iconicity along with other syntactic factors has influenced the coalescence of *have* and *to*-infinitive which initiated semantic change.

The evolution of *need* starts in Old English with *(ge)neodan* as a regular verb which developed into a personal verb by ME (see Loureiro-Porto 2009; Taeymans 2004). It was in the sixteenth century that it became used with a modal meaning. Warner (1993: 203) comments that "it develops the preterite-present form *need* alongside *needs*, and instead of disappearing, the plain infinitive continues in late Middle English […] but it is the commoner in Shakespeare, except that infinitive form *need* itself strongly prefers *to*." According to more recent contributions such as Leech et al. (2009: 94), the synchronic evolution of semi-modal *need to* appears to be independent from modal *need* and is to be interpreted as the outcome of grammaticalization of the transitive verb which semantically might correspond to a construction with the noun, e.g. *a need to do something*. Although not outspoken by the authors, this might involve syntactic analogy contrasting previous accounts, which have

claimed a movement from auxiliary-like to verb-like morphosyntax (see Taeymans 2004). Such a trajectory of development would therefore allow the occurrence of the semi-modal beyond non-assertive contexts. By contrast, Loureiro-Porto (2009: 216 f.) identifies a link between earlier uses of *need* + bare infinitive and increase in discourse frequency of semi-modal *need to* in present-day English, which she refers as retraction.

The diachronic emergence of these markers has led to the layering of the English modality system. Variation within the system reflects the degree of grammaticalization of these items in a particular time frame (Tagliamonte 2004: 34). In spite of the variable layering in the system, Myhill (1996: 341) comments that the grammaticalization of the "new" items did not replace the older ones but that these appear "to be more compatible with the general semantic organization of the language than the "old" one was." According to Myhill (1996) it is essential to focus on this particular kind of development within a linguistic category. To this end, the subjective vs. objective dichotomy within the domain of strong obligation/necessity in New Englishes will be associated to a separate process, i.e. subjectification, which operates alongside the root vs. epistemic bias (Traugott 1995: 47f).

4.4.2 Discourse-pragmatic and sociolinguistic/socio-cultural processes

Two discourse-pragmatic processes of change namely 'Americanization' and 'Colloquialization' are assumed to shape the development of present-day English. While Americanization refers to a sociolinguistic/socio-cultural or contact-linguistic process, 'colloquialization' is a more recent concept introduced by Mair (1997; 1998) to explain the contrasting developments undergone by BrE and AmE. For example, in Mair's (1998: 148-149) corpus-based study on diachronic regional contrasts the use of *help* followed by bare infinitives is an American tendency showing a marked increase in the AmE data already from 1961, whereas in the British data such a development happened later. The reverse tendency would be "Britishisation", such as for example the use of *(have) got to* at the expense of *must*, which suggests convergence of AmE with BrE norms.

Conversely, 'colloquialization' is associated with functionally motivated changes in discourse also described as "stylistic drift" towards informal and colloquial style of written English (Biber & Finegan 1989). According to Leech et al. (2009: 49), this process is primarily caused by social factors and less by linguistically motivated change. By contrast, Leech (2003) considers the two processes as inseparable. Overall, the outcome of such shifts in the conventions of formal written and spoken discourse may be correlated with general trends of democratisation processes at societal level. As contemporary societies are involved in democratising their internal structures, the less marked are overt inequalities of power at linguistic level (cf. Fairclough 1992 in Mair 1997: 204).

Within obligation and necessity, democratization involves changes in the modal meanings, such as the case of so-called "group-oriented" (intersubjective in terms of Traugott & Dasher 2002: 115) modals like *must*, *ought to* and *will*, which were replaced by more personal *got to*, 'weaker' *should* and colloquial *gonna* (see Myhill 1995; 1996). Likewise, root modality with *have to* is a more desirable option to express informal, neutral and habitual necessity that downplays overt authority and power relationships as compared with the modal *must*, which is prototypically related to direct speaker authority or even to an irresistible force of obligation (Smith 2003: 259; Sweetser 1990: 54).

4.4.3 Other determinants: the contribution of substrate and prescriptivism

Mair (2009b) suggests that the variability within the modality system in standard educated JamE is related to transfer from the creole substrate. It is therefore useful to take note of the JC modality system, which is formally similar to the one in Standard English,[52] but which differs significantly in the semantic distribution. For example, Christie (1991) analysed the two groups of verbs as exemplified in (25) which Bailey (1966: 45) identified as part of the JC modality system:

[52] Winford (1993: 85) points out the role of decreolisation in maintaining the similarities between the modality system of JC and Standard English.

(25)

Mod. 1		*Mod. 2*	
kuda	could	*mos*	must
maita	may, might	*fi*	ought to
mosa	must	*hafi*	must
shuda	should	*kyan*	can
wi	will		
wuda	should		

The differences between JC and Standard English consist in the linguistic criteria according to which the verbs assign modality, such as the positional variability of the negator and tense particle, and that these verbs do not exhibit regularity of morphological characteristics (Christie 1991: 224).[53] In her analysis on the semantics of obligation and necessity in JC, Christie (1991: 225f) observes that some of the verbs are restricted to convey only one type of modal meaning, e.g. *mosi* – probably deriving from English 'must be' –is used with epistemic reading, whereas *fi* with deontic reading (see instance (26) from Christie 1991: 227):

(26) Mi no si ar agen; shi *mosi* gaan.
I don't see her any more; she must have left.

Although the other verbs within this class – *mos*, *hafi* and *bounfi* – are considered to convey both epistemic and deontic meaning on the cline from strong to weak obligation/ possibility, it is striking that *hafi* expresses mainly deontic readings (see instances (27) and (28) from Christie 1991: 227):

(27) a. Yu *mos* did kom de kom si im. (epistemic)
You must have seen him when you came there.
b. Yu *mos* kom iina di kyar. (deontic)
You should get in the car.

(28) a. Miss Boothe *hafi* rich huom bai dis. (epistemic)
Miss Boothe must certainly have got home by this time.
b. Yu get yu dischaaj an yu *hafi* go huom. (deontic)
You've got your discharge and you have to go home.

[53] Although the JC structural composition of auxiliary elements in the verb phrase is widely accepted among scholars, it should be noted that they do not agree unanimously on how the ordering of elements follows, nor have these been studied in detail. According to Winford (1993: 88-91), Bailey's (1966) model (Modal) (Tense) (Aspect) V is ungrammatical for not being able to account for exceptions, and suggests other possible orderings. For a recent examination of the structural similarities and differences of the auxiliary ordering in JC as compared to the AmE fix system, see Milson-Whyte (2006).

Concerning the corresponding verb for *hafi* in Standard English, there are divergent opinions among scholars. According to Bailey (1966: 45), it derives from Standard English *must*, whereas Winford (1993: 93) claims that it corresponds to English *have to*.[54] To sum up, we may consider the potential influence of JC in data from ICE-JA particularly in the light of the frequency of *have to* both in private and public conversations (see Table 5.1).

As for IrE, few scholarly accounts are available on the IrE modality system. Notably, an early mentioning of the divergent use of modals in IrE (e.g. *will* and *shall*, *may* and *might*) is found in Joyce (1910: 74-77, 84). In the same contribution the use of some semi-modals such as, e.g. *'I be to do it'*, *'I have to do it'*, *'I am bound to do it'* but also *maun* ('must be') as in e.g. *'What bees to be maun be'* are emphasised as typical for the region of Ulster (Joyce 1910: 87). Thus, the modals in IrE may share features particularly with spoken Scottish English or 'Broad Scots' (Miller 1993: 99). Such typological affinity must be seen in the light of the historical circumstances of the various Early Modern English and Scots vernaculars spoken in the seventeenth century, which were brought to Ireland during the plantation settlements (Corrigan 2000: 27; Harris 1993: 141). This is in line with the complex composition of the linguistic landscape in Ireland, which cuts across the north-south boundary comprising on the one hand, northern and southern IrE, and, on the other hand, Ulster Scots with other IrE dialects (Harris 1984: 115).[55]

Additional support for non-standard influence is found in several corpus-based studies on the modality system in non-standard British and Northern-Irish dialects (e.g. Corrigan 2000; Tagliamonte 2004; Tagliamonte & Smith 2006) which have shown that the use of modal obligation/necessity in these areas differs widely from standard English. The conclusion from these studies is that modal *must* is restricted in meaning decreasing across speech groups in the whole community, while obligation with semi-

[54] A similar interpretation of modality markers is offered for Antiguan Creole in Shepherd (1982: 318). By contrast, in the analysis on the development of Guyanese Creole speakers from mesolect to acrolect, Bickerton (1975: 122-124) argues that the acquisition and use of standard English *have (to)/had* should be linked with the basilect possessive form *gat* and obligation *ga(t) fi/fu,* and which alternates with *ga(t) tu*. Likewise, the use of modal *have to* alternates with *gat to,* and which appears to be conditioned by the loss of the stative – non-stative distinction in the verb displayed in stem-form showing no tense information (cf. Bickerton 1975: 125). On the other hand, social stigmatisation and marginalisation of both possessive *gat* and *have* has led to their total avoidance in speech (Bickerton 1975: 124f).

[55] In view of the common ground shared by IrE and British dialects, Filppula (1993: 23-25) accommodates the concept of *Sprachbund* as a 'linguistic area' which comprises the British Isles, and argues that adstratal relationships have shaped the Irish linguistic contact situation.

modal *have to* appears to be the preferred variant (see section 4.4.2). Likewise, Miller & Brown (1982: 8-11) and Miller (1993: 116-119; 2004: 52) point out that in Scots modal *must* has the meaning of epistemic modality as in, e.g. *You **must** be exhausted*, whereas obligation is expressed with *have to, have got to* and *need to*.

In a more recent study on modal verbs in English and Irish, Hickey (2009) argues that the two languages lack structural equivalence, and therefore it is unlikely that modal usage displays structural transfer. The author concludes that, more than a case of substrate influence, patterns of modal usage in IrE are rather a case of analogical extension (e.g. epistemic *musn't*) through dialect input from Scots grammar (Hickey 2009: 271-272).

On the other hand, since its establishment in the nineteenth century the Irish schooling system has yielded to promote Standard English as the target.[56] Likewise, there is a long established tradition in ex-colonial English-speaking countries, such as India and Jamaica, to promote the British English model of written and spoken use. In the case of Jamaica, AmE as the exonormative norm has grown to steadily replace BrE (see section 3.2.3). However, both standards continue to be seen as prestige models (Mair 1992).

In spite of the differences in phonology and vocabulary in IndE, syntax largely follows BrE norms (Sailaja 2009: 39). As a second-language variety, a prescriptive standard seems to prevail in IndE and which can be identified in a so-called 'cline of proficiency' (Kachru 1994). As regards the standard – non-standard debate, which often aims to invalidate IndE as an established variety, Sailaja (2009: 39 f.) underlines the importance of considering level of formality and style. The case of modals and their tendency towards formal uses (Leitner 1991) could thus be the outcome of imposed prescriptive correctness.[57] According to Sailaja (2009), although the influence from AmE has been increasing more recently, for example in the media, BrE norms appear to be holding preference among educated Indians. However, it is further pointed out that

[56] Despite the importance of the National School System, Odlin (1994) shows on the basis of the figures displaying the percentage of illiteracy from the 1851 Census that schooling in English had, in fact, less impact on literacy than assumed. See also Filppula et al. (2008: 153-162).

[57] Then again, the source for increased formality in IndE may derive from the substrate as reflected for example in the Hindi-Urdu verbal morphology, which distinguishes between honorific/polite and familiar forms (Kachru 2008: 93). Obviously, these are not the only substrate languages on the Indian subcontinent, therefore the difficulty to account for all such features.

"usually native varieties become the benchmark for correctness" (Sailaja 2009: 40), be it AmE or BrE.

To conclude, there seem to be two competing models – the colonial British standard vs. North-American tendencies – which are expected to shape the use of modals in the spoken JamE, IndE and IrE.

4.5 Basic distinctions: modality, modals and related terminology

The following section introduces the specific terminology on modality, in particular the one used to describe strong obligation/necessity (4.5.1, 4.5.2 and 4.5.3).

4.5.1 Types of modal meaning

Unlike tense and aspect, modality is generally defined as the grammatical category which expresses 'non-assertion' (Palmer 2003: 5) or the grammaticalization of the speaker's subjective reflection of attitudes and opinions about the factual status or the (un)truthfulness of a proposition (cf. Bybee et al. 1994: 176). As an important discourse strategy in spoken interaction modality marks the speaker's commitment or knowledge and (dis-)agreement of the factuality and actualization of a possible or necessary situation (Huddleston & Pullum 2002: 173). Scholars usually distinguish between root/non-epistemic and epistemic modality. Despite the various extensive classifications on the type of modality[58], a binary distinction is felt as more basic and systematic for an analysis of natural language data, provided the theoretical criteria are well delineated (cf. Loureiro-Porto 2009: 20). Such a distinction concords with the historical path of the evolution of root and epistemic meanings (see Sweetser 1990) and can be easily incorporated in the study of ongoing language change. For the sake of clarity,

[58] Most often the term deontic is used to mean root or any other type of non-epistemic modality. In other cases it is used as a sub-type of non-epistemic modality together with dynamic modality which covers ability and volition, and which is opposed to epistemic meanings (Palmer 1990 [1979]; 2001; 2003). By contrast, in Goossens' (1987) threefold classification facultative is used instead of dynamic modality. Similarly, Bybee et al. (1994) and Bybee & Pagliuca (1985) distinguish between agent-oriented (as equal with deontic), speaker-oriented and epistemic, and add to these another type of subordinating modality.

throughout the present study root modality in the field of obligation/necessity is taken as the equivalent of deontic or non-epistemic modality as defined in Coates (1983), Depraetere & Verhulst (2008) and Smith (2003), and which allows the inclusion of all the possible realizations. Therefore, any deviation in the use of these terms (e.g. deontic instead of root) refers to specific choices employed by other scholars.

If there is much agreement that epistemic readings reflect the subjectification of the speaker's beliefs indicating confidence, assumption, knowledge or (dis-)agreement about possibilities in a proposition (Coates 1983; Traugott 1989), e.g. *That must be John*, quite the opposite can be said about non-epistemic/root readings.[59] Root necessity marks the speaker's authority or commitment to an action/event in a conversation, e.g. *You must go home now* (= because mother is waiting for you).[60] In this study all instances have been limited to propositions paraphrased by *'it is obligatory/absolutely essential for;…it is important for'* for root meanings as in (29), (30), (31), and (32):

(29) So I urge members <,> honourable members that we **must** evolve a viable national consensus <,> for sound fiscal management of our economy <,> </I> (ICE-India.S1B-054.txt)

(30) <#>You **have to** ride as a team<,> so that you can have one set going to the front<,> and then you have an interchange in at the regular intervals<,>and then it's a team work<,> (ICE-JA.S1B-026.txt)

(31) But the only thing is that uhm<,> sometimes it can get really crowded in there and you **'ve got to** clear the tables and serve at the bar and everything all at the same time like (ICE-Ireland.S1A-057.txt)

(32) In the century to come <,> this isssue of <,,> population aging <,> will raise many vital and new questions which **need to** be addressed urgently <,> (ICE-India.S1B-035.txt)

or 'it is necessary that…/I confidently infer that…' for epistemic meanings as in (33), (34) and (35):

[59] Coates (1983: 18-19) distinguishes between 'inferential' (i.e. involving an assumption), typically expressed with *must*, *should* and *ought*, and 'non-inferential' (i.e. involving an assessment) epistemic modality conveyed with *may*, *might*, *could* and *will*. These distinctions are placed on two ends on a scale, with confidence and doubt at each extreme.

[60] Another way to distinguish the two types of meaning is in terms of modal scope, which in case of root modality is assigned over the verb, whereas with epistemic readings over the whole proposition (on wide scope within root modality with ought to, see Nordlinger & Traugott 1997).

(33) But still <w> they're </w> losing <,>there **must** be something wrong with them <,,> like whole team (ICE-India.S1A-014.txt)

(34) There **has to** be in the Department some file on the staff member and that certainly would be the one such file we regard as official and there has to be in the University and Campus Registry some kind of official file on the staff member (ICE-JA. S1B-079.txt)

(35) You know <,> it's always **got to** be involved in some other kind of <{3> <[3> plot or or you know <,> mystique or something (ICE-Ireland. S1B-005.txt)

Even so, the two types of modality do not distinguish clearly between prototypical and ideal cases in natural language and least so in a corpus of spoken data, as any classification is prone to inconsistency. Instead, we may best characterize both root and epistemic modality in terms of gradience.[61]

4.5.2. Modal auxiliaries vs. semi-modals

Although modality in English is conveyed linguistically by several grammatical items (e.g. verbal inflections, verbal cliticised forms, adverbs or particles), modal auxiliaries (including semi-modals) are by far the most extensively studied verbal group which assign modality not only semantically but also formally. For consistency, I will refer to 'semi-modals' (see Biber et al. 1999; Smith 2003) throughout the whole study in spite of other suggestions in the literature: e.g. 'quasi-modals' (see Coates 1983; Collins 2009a; Perkins 1983; Westney 1995), 'marginal modals' or 'semi auxiliaries' (see Quirk et al. 1985), and even 'emerging modals' (see Krug 2000).

A concise description of the present-day English modals and semi-modals is found in Quirk et al. (1985: 120-148), who define auxiliarihood as gradient involving auxiliary and main verb properties such as: (a) core modals: *can*, *could*, *may*, *might*, *shall should*, *will*, *would*, *must*; (b) peripheral or marginal modals: *dare*, *need*, *ought*; (c) modal idioms: *had better, would rather/sooner, **be** to, **have** got to* etc.; (d) semi-auxiliaries: ***have** to, **be** about to, **be** able to, **be** bound to, **be** going to, **be** obliged to, **be** supposed to, be willing to*, etc.; (e) catenatives: ***appear** to, **happen** to, **seem** to, **get** + -ed*

[61] On the concept of *gradience* in grammar, see Quirk et al. (1985: 90).

participle, ***keep*** + *-ing* participle, etc.; (f) main verb + non-finite clause: ***hope*** + *to*-infinitive, ***begin*** + *-ing* participle, etc. In view of this classification *must* belongs to group (a), *(have) got to* to group (c) and *have to* to group (d). In this study *(have) got to* is considered a semi-modal. Likewise, unlike Krug (2000), *need to* features as a semi-modal motivated by its increase in frequency in BrE and AmE (Leech et al. 2009: 94). The classification above is in accordance with the formal characteristics of modal auxiliaries also known as 'NICE' properties (Huddleston 1976 in Palmer 1990 [1979]: 4-5):

Negative contraction	*mustn't*
Inversion with subject	*Must* we go?
Coda	She *must* go, so *must* he.
Emphatic affirmation	But you *must.*
No *–s* form for third person sg.	**musts*
No non-finite forms	**to must*, **musting*
No co-occurrence with modals[62]	**will must*

In addition, further properties distinguish the core from marginal modals, for example, *may* in negation (?*mayn't*)[63] is not phonologically reduced, *can* does not possess a clitic as *will* (*'ll*) whereas *must* cannot form the preterit (Collins 2009a: 14). By contrast, semi-modals do not follow any of the NICE criteria and use *do* as an operator in inversion, negation or in emphatic affirmation.[64]

One of the main advantages of the *have to* and *need to* is their flexibility on morpho-syntactic level (Collins 2005: 256), which is different from both *must* and *(have) got to*. Thus, semi-modals use *–s* form for third person singular and can be used with non-tensed as well as with tensed forms, which, in the case of *must* and *(have) got to* is not possible. A further indicator for these differences is to be found in negations. Consider for example sentences (36) and (37) with focus on the morhposyntactic realisation:

(36) You **must** *not* be facing that problem you must be staying somewhere close (ICE-India.S1A-065.txt)

[62] On double modals in non-standard varieties of English, see Nagle (2003).
[63] Arguably, the form *mayn't* is likely to occur occasionally in BrE (Trudgill & Hannah 2002: 61).
[64] Although the auxiliary-like use *haven't to* is attested in the development of the obligation meaning of this marker (Denison 1993: 317), currently it is obsolete, and only *haven't got to* has been reported in non-standard BrE (Tagliamonte 2004: 40).

(37) In other words you *don't* **have to** be formally married to ask a maintenance right (ICE-JA.S1A-099.txt)

The negative construction *do not* + *have to* is appealing to speakers also due to the scope of negation. It also indicates common features with main verbs asserting non-performance of necessity (Smith 2003: 244). As such, negation with semi-modal *have to* in (37) seems to be less authoritative and can be paraphrased as e.g. *It is not necessary to formally marry*, which is in fact a negation of the whole modality. On the other hand negation with *must* in (36) has scope over the verb phrase as in e.g. *It is necessary for you not to be facing that problem*. Further on, *have to* contrasts with *(have) got to* in regard to the availability of past verbal forms (e.g. *And I had to* (?*had got to*) *crawl across the railway lines to get out of the way of the Dublin Express* (ICE-Ireland.S1A-001.txt)). Having both as first element an auxiliary, structurally the two constructions are very similar. In addition, as Krug (1998; 2000) claims, the emergence of the grammaticalized reduced form *gotta* is evidence for its status as an auxiliary. As for contracted forms of semi-modals *('ve to*; *'ve got to/'s got to*), they are potential indicators of style variation (for instance colloquial/informal style), and can offer valuable insights in our attempt to uncover ongoing change.

Apart from the structural distinction between modal and semi-modal, previous descriptions have focussed on the source of necessity. This will be the object of the next section (4.5.3).

4.5.3. Subjective vs. objective modality

A widely assumed distinction associated with the modal – semi-modal alternation refers to a 'subjective' function of root modality typically with modal *must* (38), and an 'objective' function with semi-modal *have to* (39) (Coates 1983; Huddleston & Pullum 2002; Leech 1987; Quirk et al. 1985):

(38) So first you **must** have a good working definition of bioterrorism (ICE-JA.S1B-005.txt) [speaker involvement]

(39) First you **have to** solve English language and then you have to solve reasoning ability (ICE-JA.S1A-071.txt) [exam regulation]

Although in a more subtle way, Depraetere & Verhulst (2008: 11) argue that the subjectivity of the root meaning with *have to* in (40) is conveyed by the speaker's opinion, by using the hedge *I think* to stress the urgency/necessity of the situation:

(40) First of all *I think* the government **has to** make its position clear (ICE-India.S1B-025.txt) [personal opinion]

The notion of subjectivity used here – i.e. not conventionalized relation between structural and semantic change – is quite different from the cognitive process discussed in section 4.4.1.1, which involves the speaker's interpretation of a proposition through grammatical change (i.e. change from root to epistemic modality).[65] Although subjectivity does not necessarily entail subjectification, these two notions will be approached as complementary phenomena as part of the dynamic process of layering and variation in synchronic data. More specifically, I suggest that the subjective element of the root reading of *have to* in (40) can be inferred pragmatically as a case of subjectification, which might account for an intermediate stage on the unidirectionality cline of change towards even more subjectified meanings based on the speaker's beliefs or knowledge (see section 6.3.3.2). Such an interpretation would, then, correspond to Traugott's (1989: 35) third tendency involved in semantic change. Notwithstanding, the relationship between subjectivity and subjectification is complex, but which can be usefully combined for the understanding of the asymmetries found in synchronic data from New Englishes.

Traditionally, the distinction between subjectivity and objectivity refers to speaker-relatedness (speaker involvement) or non-speaker-relatedness (external orientation) in a modal utterance (Verstraete 2001: 1506). Lyons (1977: 792-793) links the notion of subjectivity particularly to epistemic modality, but the same distinction is relevant to deontic modality as conveyed by the speaker's involvement – usually as the authority compelling the addressee to some activity in the utterance. In addition, Lyons (1977: 797-801) suggests that epistemic modality can be assigned both as subjective and objective.[66]

[65] Consider, by contrast, Krug's (2000: 61) observation on the role of subjectivity in the rise of obligative *have to* from its initial possessive nonmodal semantics which re-affirms that subjectification affects various stages of change in the grammar in general, and in modality more specifically.

[66] On subjectivity in epistemic modality, see Nuyts (2001). See also Verstraete's (2001: 1508) evaluation of Lyons' distinction between the interpersonal and non-interpersonal function of language.

On the basis of Lyons's (1977) semantic account, Palmer (1990 [1979]: 10-11) claims that the English modals are essentially subjective, unlike semi-modals (e.g. *have to*) which are assumed to express objectivity.[67] For this reason, Verstraete (2001: 1508) assesses deontic modality as a problematic category. Likewise, Westney (1995: 97) argues that a systematic distinction is found in the subjective vs. objective contrast within root modality, which are often linked to notions such as 'source' and 'target' of necessity. In this situation, *have to* and *(have) got to* as the newcomers in the modal system have a narrower and more specialized meaning than modal *must* (Westney 1995: 97). A somewhat different terminology is found in Biber et al. (1999: 494-495), who talk about *intrinsic* and *extrinsic* meanings, with *must* as the only modal which expresses both meanings, whereas *have to* appears to be the most common form for personal obligation in conversations. In addition, Coates (1983: 36, 55-56) claims that the contrast between subjectivity and objectivity with the two more or less interchangeable markers *must* and *have to* correlates with a range of grammatical features: e.g. second person subject; speaker has authority over subject; type of verb (agentivity); with the underlying paraphrases *'it is obligatory/absolutely essential that'*; animacy; *'it is important that'*; and inanimacy. A particular feature which differentiates the two markers is found in the source of root modality which she refers to as speaker involvement, a concept that Palmer (1990 [1979]: 36) labels as discourse-oriented:

(41) "You **must** play this ten times over", Miss Jarova would say, pointing with relentless fingers to a jumble of crotchets and quavers (Coates 1983: 34) [the speaker is the source of the obligation which is directed to the addressee → subjective modality]

(42) there is already a great imbalance between what a student **has to** pay if he's in lodgings and what he **has to** pay if he is in a hall of residence (Coates 1983: 55) [the speaker is completely neutral → objective modality]

According to Coates's parameters, example (41) is indicative for the 'core' function of modal *must* which expresses speaker involvement through the second person animate subject, the activity verb, and can be paraphrased as *'it is important that'*. The semantics of *have to* in (42) is still of root modality, however the speaker represents only the 'channel' who communicates the necessity (Depraetere & Verhulst 2008: 3) as in the

[67] See Tregidgo (1982: 81) for a similar assertion.

paraphrase *'it is necessary for'*. Thus, the authority is not overtly expressed.[68] For Depraetere & Verhulst (2008: 3) the distinction between subjectivity and objectivity usually points to some contextually present reference as its source, i.e. "the driving-force behind the necessity", and which can be either the speaker or some other external entity. This, in turn, ties in with Jespersen's (1924: 320) assertion that obligation/necessity contains an 'element of will'. Two major types of sources, i.e. discourse-internal and discourse-external, emerge from their classification (see Table 4.1):

Discourse-external sources (objective)	**Discourse-internal sources (subjective)**
Rules and regulations	The speaker/hearer in statements
Circumstances	Personal opinion (*I think, I believe, in my view*)
Conditions	

Table 4.3. Discourse-internal vs. discourse-external sources of necessity (adapted from Depraetere & Verhulst 2008: 4)[69]

The corpus analysis combines this classification with Coates' (1983) parameters which, according to Tagliamonte (2004: 43), function as "grammatical diagnostics".

In sum, the discussion about the functional orientation of necessity is closely linked with semantic reading. Table 4.2 presents the relationship between subjectivity and subjectification, as a mechanism of semantic change, and the intermediate/transitory uses within root modality:

[68] Apart from the assumed restriction of semi-modals to acquire subjective root meaning, Westney (1995: 55-56) rightly questions the claim about their non-availability of either subjective or objective epistemic meaning in corpus material. Witness also Sweetser's (1990: 65 f) claim that an interpretation of such contrasts within root modality is conditioned by pragmatic factors. See also Coates (1983: 15-17) on polysemy between root and epistemic modality in cases of indeterminate examples which she includes in a separate category of 'merger'.

[69] In addition to these two types of sources, Depraetere & Verhulst (2008) report mixed sources, a combination of both discourse-internal (i.e. personal opinion) and discourse-external (i.e. specific circumstance or a condition) necessity. However, in the present study I opted to emphasise the most overt source involved in the necessity.

		Subjectification →		
Root necessity			*Epistemic necessity*	
Objective ⇢	Subjective		Objective →	Subjective
Discourse-external	*Discourse-internal*		*Non-inferential* (i.e. logical necessity)	*Inferential* (i.e. confidence)
(e.g. Depraetere & Verhulst 2008; Goossens 2000; Traugott 1989; Traugott & Dasher 2002)			(e.g. Coates 1983; Lyons 1977)	

Table 4.4 The relationship between source of necessity and subjectification

As shown with *must* and *have to*, this distinction also occurs with the other two expressions *(have) got to* and *need to* as in (43), (44) and (45) from ICE:

(43) Uh now <,> I **'ve got to** work this out (ICE-GB.S1A-070.txt) [the speaker as the source of necessity]

Coates (1983: 57) claims that *(have) got to* marks subjective necessity contrasting it with objective *have to* and associates it semantically with *must*. Likewise, recent studies on long-term diachronic change in BrE and AmE (see Krug 1998a; 2000; Myhill 1995; 1996) emphasise the strong subjective force associated with this expression. Considering such claims, the source of necessity of the contracted form *'ve got to* in (43) appears to be internal to the speaker suggesting an instance of 'self-exhortation' as exemplified in (1).

Subjectivity and objectivity are useful terms to explain the semantic differences between modal *need* and semi-modal *need to*. Thus, the modal expresses external/objective force, whereas the semi-modal traditionally conveys subjective (personal) wish as internally-motivated necessity, but lacks the force of a personal directive (Nokkonen 2006: 37). However, instance (44) contradicts such assumptions on *need to* as it may express also speaker-imposition, whereas (45) suggests that this semi-modal can also be associated with objective sources:

(44) But you're a child ma'am <#>This man has made you into a child <#>You **need to** make up your mind whether you want to be a grown person or somebody's child because only my child I can tell not to go to the gate (ICE-JA.S1A-099.txt) [speaker gives a recommendation → subjective]

(45) Those jobs don't really fit here and we **need to** look at the package and see the package in this respect (ICE-JA.S1B-073.txt) [external source of necessity → objective]

The subjectivity of *need to* in (44) is rendered by the pragmatic force with which the speaker imposes the obligation, and which in this case appears to be a recommendation for the addressee's own sake. For Smith (2003: 245; 260), such uses which do not express overt objective obligation increase the ambiguity of this semi-modal, which is reflected here in the speaker's concern for the listener's needs. Then again, the recommendation seems to hide an order or instruction. At the same time, if the semi-modal would be replaced with *must* one would not note anything odd in the meaning of the utterance, apart perhaps from the assumption that the necessity conveyed with the modal is more intense.[70] By contrast, (45) is not speaker-related as the necessity lies outside the imposition of the speaker (i.e. *those jobs* as the source of necessity), and is thus objectively motivated.

4.6 Limitations of the present study

At the end of this chapter, I should emphasise some of the limitations of the whole study. As already mentioned in the previous sections, the present study is intended to focus only on one domain of English modality, namely the way strong obligation and necessity is used in spontaneous interaction from New Englishes. Given that spoken data often contains dysfluencies (Myhill 1996: 346), the transcription of the dialogues brings into discussion the problem of potential inaccuracies, even if these lie within reasonable margins of error. For example, a problematic case is the inclusion of the full and contracted variants of the semi-modals, such as *'ve to* for *have to* or *'ve got to* and *gotta* for *have got to*, and which, if not considered, might skew the overall findings (Krug 2000: 38).

A further point relates with the issue of indeterminacy of modal meanings. As corpus data may contain a high degree of indeterminacy and, hence, ambiguity, it is often the case that the researcher has to decide for one meaning over the other. To some

[70] See also the discussion on the alternation between *must* and *have (got) to* in Palmer (1990 [1979]: 115).

degree root and epistemic meanings are overlapping. Certainly, the linguistic context plays an essential role in this case. Even though each instance of *must*, *have to*, *(have) got to* and *need to* was assessed several times according to the paraphrasing method and the wider discourse context, a strict delimitation of semantic contrast was often difficult to establish. It is therefore hoped that the criteria used for assigning of either root or epistemic meanings are reliable for a sound and objective classification, and that any inconsistencies are minimal so that they will not affect the overall outcome of the investigation.

As regards the modal – semi-modal alternation, it will suffice to point out the main descriptive problems so far in the literature and the way this situation can be remedied with respect to the New Englishes. The systematic evaluation of such distinctions and their use in spoken data will offer a solution to the previous descriptive problems applicable to the present data set. In how far trends based on the New Englishes investigated here can be generalized further remains to be seen.

Further important issues refer to interpreting the statistics and how to relate the synchronic form-function relationship with diachronic developments of modal obligation/ necessity in the New Englishes. Despite consistent evidence from the standard and non-standard Englishes, in the absence of diachronic data from the three New Englishes under study the decision whether the uses of a variant form indicate 'colonial lag', are a local innovation via substrate transfer, learner strategies or are part of "directional drift" (Dollinger 2008: 146-148, 168-170) remain tentative at this stage. Nevertheless, the special issue of the Journal of English Linguistics on the history of modality in the New Englishes edited by Noël et al. (2014) shows that this research gap is beginning to be filled at least for the cases of AusE, PhilE, WAfE and BSAfE.

5 Competing variants in ICE: description of trends

5.1 Introduction

Having outlined the major theoretical and methodological aspects of the present study, I will now turn my attention to analysing the similarities and differences of formal patterns of obligation/necessity in data from ICE-JA, ICE-India and ICE-Ireland. Although variation between *must*, *have to*, *(have) got to* and *need to* involves particularly semantic contrast, our interest in this chapter is centred on form. Assuming that there is a link between synchronic regional and stylistic variation and (long-term) change (Collins 2009b: 283), with discourse frequency functioning as a determinant in such processes, the asymmetries in the distribution of forms in the data suggest a different language dynamic in standard JamE, IndE and IrE. Some of the formal characteristics of these markers in ICE, and particularly in ICE-Jamaica, have already been examined in Mair (2009b). Against the backdrop of these findings, I will extend the discussion of the modal – semi-modal alternation from a cross-variety perspective in order to reconstruct the successive stages of ongoing change from the synchronic distributional constructs.

The purpose of this survey is to quantitatively assess the synchronic distribution and the relations between the four modal items by way of descriptive statistics, which will serve as background for the more detailed discussion of their orientation towards a British, an American or a local norm. What this survey aims to explore, then, is whether: (i) considering the specific socio-historical contact situations in the emergence of JamE and IndE the distribution of forms in the corpora of these two varieties lends support to an overt 'conservative' profile, as shown by the consistent use of the modal *must*; (ii) and whether IrE, as a new L1 having started out as a language-shift variety, follows largely British usage patterns. Such an approach becomes more interesting as these four modal items are often analysed as near semantic equivalents in present-day English, but which are associated to different levels of formality. Thus, such variability entails that, unlike their modal counterparts, semi-modals are commonly associated with informal/colloquial style.

Since the spoken medium is more susceptible to change in terms of fluctuations in frequency of use the quantitative assessment might capture different stages of

development within this grammatical area in these three varieties (cf. Krug 2003: 10). Therefore, attention will be paid to the behaviour of these forms in spoken (private and public) dialogues in each of the subcorpora in ICE. To determine whether their layering in New Englishes patterns more with a native-like profile, stylistic variability will be correlated to level of formality (for the specific case of JamE cf. Mair 2009b: 20). To this end, I will concentrate on medium-specific (i.e. spoken dialogues vs. written press texts) and formality (i.e. formal vs. informal contexts of speech production) differences which provide insights about the linguistic choice in spoken interaction. An additional examination of the syntactic and stylistic characteristics of this markers will help identify the extent of variation.

The data under analysis comprises a total number of 6142 tokens representing both spoken and written texts from ICE-JA, ICE-India and ICE-Ireland. In order to provide a balance between the inner and outer circle, data from SBC and Frown have been added to represent spoken and written AmE. As the specific corpus design of SBC and Frown does not allow a straightforward comparison with ICE, frequency of occurrence was normalized per 100,000 words. Using normalized frequency counts as a discriminating measure potentially increases the explanatory power of synchronic findings, whether the distribution of items across spoken texts (and later across written texts) is stable or whether innovative uses emerge (cf. Krug 2000: 86).

At the end of the chapter, the main regional and stylistic trends will be summarized from the perspective of the degree of convergence with and/or divergence from standard international norms of English as conservatism (also ‘colonial lag’) or local independent developments (see 2.2).

5.2 Distribution of forms in the spoken section

Due to the same number of dialogues comprised in each category (private and public), as well as the same type of conversations, data from ICE lends support to some generalizations about synchronic contrasts in the expression of strong obligation/necessity in New Englishes. Figure 5.1 displays how each variety patterns

with each variant form (for a detailed view of all raw and normalized frequencies, see Appendix 3.a).

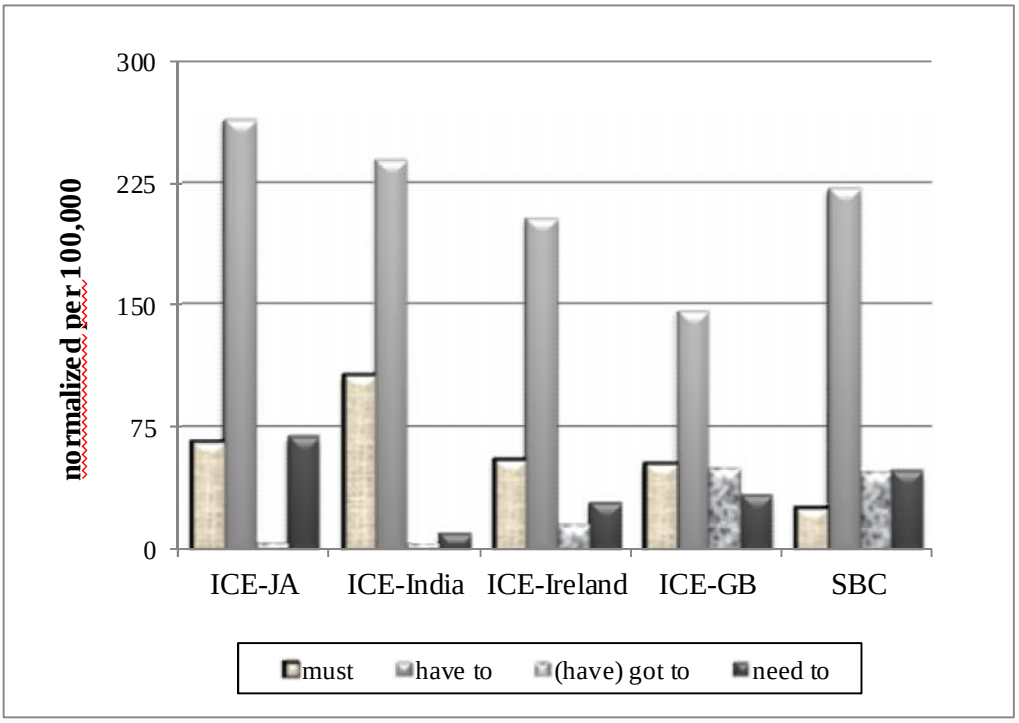

Figure 5.1. Overall distribution per 100,000 words of *must*, *have to* and *(have) got to* and *need to* in the spoken dialogues.[71]

At this first stage of the analysis, overall, the regional and stylistic distribution correspond with the figures presented in Mair (2009b). Significance of variation between inner and outer circle was statistically assessed on the basis of Chi-square test (Oakes 1998: 26). The overall distribution of the four modality markers in the three regional datasets – ICE-JA, ICE-India and ICE-Ireland – was tested for significance in turn as correlated with ICE-GB and SBC. The values obtained indicate very high significance at $p<.000$,[72] which confirms that the observed differences between the three focus varieties and BrE as well as AmE are likely not due to chance. This evidence is taken as starting point for the more detailed form-based analysis.

[71] This graphic considers all the morphological variants *have to*, *has to*, *having to*, *had to* as well as contracted forms *'ve got to/'s got to*/*gotta*; *need to*, *needs to*, *needing to*, *needed to* and negated forms.

[72] ICE-JA: ICE-GB χ^2 = 242.81, d.f. 3, $p<.000$; ICE-India: ICE-GB χ^2 = 305, d.f. 3, $p<.000$; ICE-Ireland: ICE-GB χ^2 = 98.2, d.f. 3, $p<.000$; ICE-JA: SBC χ^2 = 188.15, d.f. 3, $p<.000$; ICE-India: SBC χ^2 = 346, d.f. 3, $p<.000$; ICE-Ireland: SBC χ^2 = 89.8, d.f. 3, $p<.000$.

The most striking differences are found in the variable ranking of modal *must*, semi-modals *(have) got to* and *need to*, whereas *have to* appears to function as the default variant form in all five varieties. Especially the Indian and Jamaican data display high preference for semi-modal *have to* which contrasts with a lower frequency rate of *must*. Such distribution seems to lend support to the weakened status of the modal in present-day English.

Looking separately at the distribution of modal *must*, the most prominent variety is IndE (107.7 occurrences per 100,000 words) with an incidence twice as high as in ICE-GB (53 occurrences per 100,000 words), whereas SBC (26.1 occurrences per 100,000 words) takes the lowest position within this arrangement. Such a result is not surprising considering the recent trends which point towards a decline in the frequency of this marker especially in AmE (Leech 2003; Mair & Leech 2006; Smith 2003). Thus, the data form SBC is most suggestive for such tendencies. The high frequency in both ICE-India and ICE-JA confirms the initial assumption in regards to the consistent use of modal *must* in the outer circle. Moreover, the data show that JamE (with 66.1 per 100,000 words) and IrE (55.8 occurrences) take, so far, an intermediate position within the five world Englishes under study.

On the other hand, the lower distribution in ICE-JA as opposed to ICE-India seems to suggest that the use of this modal in JamE is moving along recent global trends of decrease in frequency, even though at a slower pace than in the native varieties of English. By contrast, the very high frequency of this marker in IndE possibly denotes 'colonial lag' in the form of an increased awareness to use the prescriptive norm. One might be tempted to infer that the use of strong obligation and necessity in IndE and BrE differs only with regard to ranking in their distribution, an assumption already expressed in Leitner (1991). Following this line of reasoning, a comparison of Leitner's findings for modal *must* with those provided in Leech (2003) from the Brown corpora shows few deviations in their distribution indicating that written IndE largely follows native British norms. However, arguing that spoken language is equal to the written norm in IndE is only tentative at this stage. Note, that the conservative character of IndE as concerns the uses of this modal has been stressed in several recent studies (e.g. Collins 2009b; Mair 2009b).

In addition, the cross-variety comparison shows that IrE follows a British pattern of use as the data shows almost full equivalence in distribution with ICE-GB. Overall, although ICE-India as the only variety standing out the line, the comparison of frequencies with modal *must* reveals that ICE-JA, ICE-Ireland and ICE-GB provide a fairly balanced distribution. Conversely, in SBC this modal is on its lowest position in the system. From this one-dimensional perspective AmE seems to be the most advanced variety. However, a direct link to either British or American patterns of use is still difficult to establish at this stage.

The main distinctions occur in the distribution of semi-modal *have to*, which by far outnumbers the occurrences of *must*, indicating that the latter is presumably less preferred in spontaneous interaction. Regardless of whether inner or outer circle, all the varieties display very high frequency counts with this semi-modal, however, in different proportions. Although Figure 5.1 shows that, overall, BrE, IrE and AmE make use of a wider range of semi-modals, the incidence of *have to* is highest in JamE and IndE. Thus, both the Jamaican (with 264.7 occurrences per 100,000 words) and Indian (with 240 occurrences per 100,000 words) collection of conversations are in the leading positions, with JamE displaying the highest frequency altogether. Since the difference between ICE-JA and ICE-India is relatively small, amounting to a disparity of roughly 25 instances per 100,000 words, it can be assumed that the two varieties pattern alike. This is somewhat surprising, as Collins's (2009b) study on modality in World Englishes shows that this semi-modal occurs by far most often in AmE.[73] Conversely, BrE (with 146.9 occurrences per 100,000 words) appears to be lagging behind in this arrangement. While the frequency of *have to* in AmE (with 240 occurrences per 100,000 words) is here in the highest ranking within the inner circle varieties, IrE (with 203.8 occurrences per 100,000 words) ranges even higher than BrE.

Considering the alleged functional equivalence between *must* and *have to*, the high discourse frequency of the semi-modal seems to indicate that it is taking over the meanings of its modal counterpart in the two outer circle varieties more than in those from the inner circle. However, establishing whether the widespread use of *have to* involves generalization of meanings in the two outer circle varieties is an issue which remains to be examined in the next chapter (see chapter 6).

[73] Note, however, that Collins (2009b) establishes this ranking on the basis of both spoken and written data from ICE.

A striking finding is that the occurrences with semi-modal *(have) got to* are quite low in our data, particularly in ICE-JA and ICE-India. In fact, the most pronounced disparity is found in the distribution between semi-modal *have to* and *(have) got to* in ICE-JA and ICE-India, which yields no comparison with the three native varieties. Irrespective of their common structural features which the two constructions share to a certain extent, they followed distinct paths of evolution (see 4.3.1.3) as the frequency of *(have) got to* has increased because of the fast development of the variant forms *got to* and *gotta* in present-day English.[74] This does not seem to be the case in ICE-India and ICE-JA, and is consistent with findings reported in Collins (2009b: 288-289) on other world Englishes and van Rooy & Rossouw (2011) on SAfE, but contrast with the Antipodean varieties of English (e.g. AusE and NZE) as described in Collins (2005: 253). Contrary to the lack of evidence from non-standard Northern IrE dialects (Corrigan 2000: 35), *(have) got to* is present in ICE-Ireland (with 15.8 occurrences), yet less than half the amount found in ICE-GB (50.8 occurrences) and SBC (47.7 occurrences). In spite of the popularity of *(have) got to* of present-day English, its low frequency counts in the dialogues of the three British derived varieties is noteworthy, and emphasises once again the problematic concerning the different linguistic ecologies in these varieties. The fact that this marker is a recent development in English is a possible reason for its low usage in many British-derived colonial Englishes.

A recent innovation gaining ground in present-day English is semi-modal *need to*, which is assumed to enter competition with *must* and *have to* (see Nokkonen 2006; Smith 2003). In contrast with the fairly even distribution of *must*, *have to* and *(have) got to* in the outer circle varieties, the frequency counts for *need to* are disproportionate. As an exponent of internal necessity, *need to* neutralises the effect of overt subjective obligation and is therefore perceived as more polite than modal *must*. Tentative support for the hypothesis that *need to* has entered competition with the other markers of necessity is available for example in ICE-JA, where modal *must* is slightly less frequent than this semi-modal. Such distinctions are worth examining in more detail in ICE-JA. Moreover, ICE-JA is leading the way with an incidence of 69.4 occurrences per 100,000 words followed by SBC with 49.3 occurrences, whereas ICE-Ireland and ICE-GB show an almost equal distribution of 29.4 and 33.8 occurrences, respectively. In

[74] Consider Krug's (1998a: 179; 2000: 63) claim with regard to the rapid increase of variant form *gotta* in BrE and AmE which he named as a "success story" in the English syntax of the last 150 years.

stark contrast with ICE-JA, ICE-India displays the lowest number of this semi-modal, with an incidence of 10.5.

Despite inherent differences, overall, the data display regularity and, to some degree, analogy in the frequency rank – as we will see also in the comparison between private and public dialogues in Figure 5.3 and 5.4 – both in the inner and outer circle. Table 5.1 below summarizes the main synchronic trends in the order of their occurrence in the present data from the perspective of layering of competing variants:

Variety	**Rank of ordering**
ICE-JA	*(have) got to - must - need to - have to*
ICE-India	*(have) got to - need to - must -have to*
ICE-Ireland	*(have) got to - need to - must - have to*
ICE-GB	*need to - must - (have) got to - have to*
SBC	*must - (have) got to - need to - have to*

Table 5.1. Variable ordering of strong obligation/necessity in spoken texts

The ranking in the table above is frequency-based and should not be confused with a representation of the diachronic evolution. The table shows that *have to* outnumbers in regard to its frequency of use the other markers of necessity in all the varieties under study. The lowest positions in this ordering are occupied by *(have) got to*, *need to* and *must*.

By far, the most interesting result is that variation in spoken data from JamE, IndE and IrE is most robust between *must* and *have to*, whereas the use of *(have) got to* turns out as a marginal variant in the outer circle and *need to* is restricted to high preference in ICE-JA. Two opposing forces seem to be at work in the data from ICE-India. While the high acceptance of modal *must* and very low incidence of the more recent expressions *(have) got to* and *need to* suggest conservative status, researchers on non-standard varieties of BrE have interpreted the overwhelming preference for *have to* as a case of "ongoing retention of an older layer in the development of forms" (Tagliamonte & Smith 2006: 356). Based on such reasoning, it might be suggested that the Indian and Jamaican data reveal a higher level of formality in speech which maintained its position in the post-colonial period. Conversely, the low rate of

frequency of *(have) got to* in the Irish data is less striking if we analyse it against lack of use in non-standard data from South Armagh English (Corrigan 2000: 35) or even absence of *must* in the dialects spoken in Northern Ireland (Tagliamonte & Smith 2006: 357).[75] The complexity of such interrelations adds more puzzles to the arrangement of items within the modality of world Englishes.

Although the specific differences between BrE and AmE are not the focus in the present study, the frequency patterns obtained so far are indicative for the most recent trends which these constructions have followed in present-day English. The logical conclusion from this frequency-based ordering in the spoken dialogues appears to be that BrE is lagging behind not only AmE but also the other three varieties. Notable in this sense is the use of *(have) got to* as a recent British innovation, which is paralleled by *need to* in AmE.

Of relevance for the synchronic evaluation of frequency patterns with semi-modal *have to* are the various stages of its development in present-day spoken English. Krug's (2000: 88) study on variation between *have to* and *(have) got to* in spoken interaction from BNC indicates that frequency of use with *have to* reaches its peak in present-day BrE in the 90's with adult speakers aged 45-59, whereas the lower use among younger speakers points towards a stabilization of this semi-modal amounting to 1,700 occurrences per million words. As for present-day spoken AmE, evidence for its frequency of use is provided in Biber et al. (1999: 488) indicating a slightly higher incidence per million words than in BrE. In view of such findings, the contrasts in the distribution of *have to* particularly in ICE-JA and ICE-India would indicate that this semi-modal might not have reached saturation yet, and that the corpus analysis captures an earlier stage of development of this marker in the outer circle.

Similarly, the critical period for the increase of discourse frequency of *(have) got to* is considered the early twentieth century. From a cross-variety perspective, the most likely explanation for the low spread of this expression is that a recent syntactic Briticism of informal nature did not make it into these varieties.[76] Incidentally,

[75] The comparison between the ICE data and dialect data in Tagliamonte & Smith (2006) is quite limited. However, the evidence of ongoing change as shown in dialect data from the British Isles is an important source of information as regards conservatism or innovation reflected in other standard varieties of English.

[76] See section 4.4.1.3 with regard to Krug's (2000: 64) claim on cognitive processing constraints as a barrier in the course of conventionalization of utterance.

geographic proximity between Jamaica and North America does not seem to contribute to increased usage of this expression. While under these circumstances the spoken acrolect variety in Jamaica appears to fulfil rather the functions of the written standard norm (see Shields-Brodber 1997), further parallels can be drawn to the creole substrate. Unlike the JC modality marker *hafi*, which potentially increases the use of *have to* in the acrolect, there is no corresponding basilect form for informal *(have) got to* (as regards stigmatization in creole use, see also footnote 54). On the basis of such evidence, it seems reasonable to assume that the function usually expressed with *(have) got to* has been taken over by one of the other three markers under study.

The diffusion of *must*, *have to*, *(have) got to* and *need to* in the spoken dialogues capture not only the different stages of diachronic development but also a possible specialization of *have to* and *must* in the New Englishes.

5.2.1 Range of variation across formality level and certain text types

If section 5.2. has provided a general overview of the frequency patterns in the spoken dialogues in the outer and inner circle varieties under study, it seems appropriate to take a further step and examine the forms across text categories of both formal and informal contexts. While Figure 5.1 is indicative for the general trends in the use of modal obligation/necessity in the four varieties, the present section sheds some more light on the stylistic preferences in the formal and informal contexts from the spoken conversations.

For several obvious reasons which have been already outlined (see 4.3.1 and 4.3.2), the specific design of SBC does not provide the appropriate framework to undertake such an exploration. Therefore, a quantitative analysis by text category with this corpus is not considered here. The distribution of the individual items is presented separately representing private (see Figure 5.2) and public dialogues (see Figure 5.3).

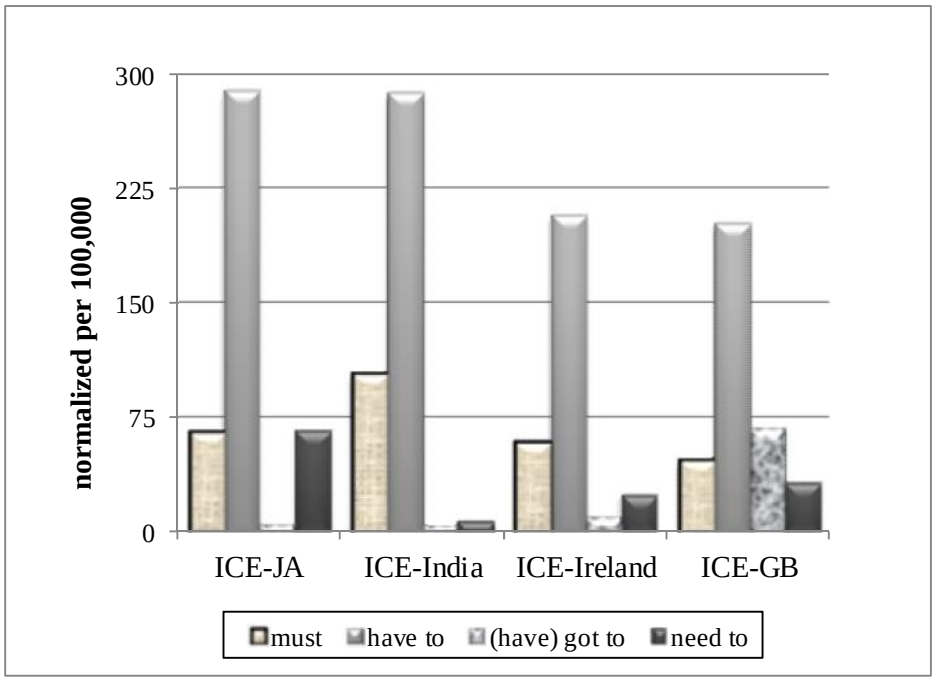

Figure 5.2. Distribution of strong obligation/necessity per 100,000 words in private dialogues (text files S1A-001 to S1A-100)

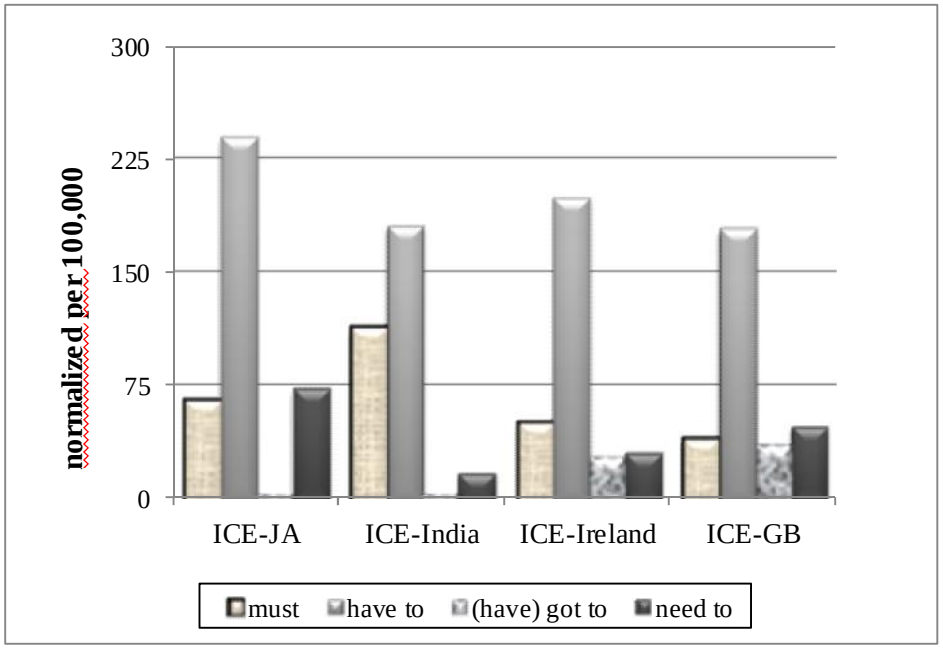

Figure 5.3. Distribution of strong obligation/necessity per 100,000 words in public dialogues (text files S1B-001 to S1B-080)

Looking at the overall distribution of forms in private dialogues, it is striking that the frequency within the outer and inner circle taken separately seems to be quite balanced. Such an assumption is validated by the Chi-square test which has yielded very high

significance at the level of $p<.0001$.[77] Moreover, Figure 5.2 shows similarities with the patterns found in Figure 5.1 (excluding SBC). Thus, ICE-JA (with 290.5 occurrences per 100,000 words) and ICE-India (with 288 occurrences) show almost full equivalence in the distribution of semi-modal *have to*, whereas ICE-India displays again a high incidence of modal *must* with 104.5 occurrences. By contrast, ICE-JA, ICE-Ireland and ICE-GB reveal lower preferences for this marker. Contrary to expectations, the semi-modal *(have) got to* shows again very low frequency of occurrence both in the two outer circle varieties and in IrE. Although it is most often associated with obligation based on emotion (Myhill 1995), speakers in these three varieties seem to avoid it consistently in spontaneous interaction. Only ICE-GB displays consistent use of the semi-modal *(have) got to* outnumbering *must* and *need to*. However, Krug's claim with regard to the steep increase of this marker in contemporary BrE cannot be validated here.

A further interesting aspect is the almost equal distribution between *must* and *need to* in ICE-JA. Moreover, the incidence of this semi-modal (69.4) appears to be highest in ICE-JA, exceeding the incidence found in ICE-Ireland and ICE-GB. Furthermore, *need to* in ICE-JA is almost equally frequent as *(have) got to* in ICE-GB. By contrast, strong necessity in ICE-India and ICE-Ireland appears to be highly restricted to the use of *have to* and *must*. This is quite striking as both *must* and *need to* are associated primarily with the expression of internally motivated necessity. Given such balanced distribution it appears that layering within strong obligation/necessity in JamE informal speech functions differently from the other varieties.

Moving over to the distribution of forms in the public dialogues (see Figure 5.3) representing formal contexts of speech a different pattern is found. This is captured also in the results obtained from the Chi-square test, which yielded very high significance ($p<.0001$) for the evaluation of ICE-JA as correlated with ICE-GB and ICE-India with ICE-GB, but which yielded no significance in the evaluation of ICE-Ireland as correlated with ICE-GB, as $p<.147$ (not significant).[78] The 'non-significant' result shows that, at least as concerns the formal contexts of interaction, IrE follows the BrE pattern. Given the different text categories in the section of public dialogues, such

[77] ICE-JA: ICE-GB $\chi^2 = 153$, d.f. 3, $p<.0001$; ICE-India: ICE-GB $\chi^2 = 197$, d.f. 3, $p<.0001$; ICE- Ireland: ICE-GB $\chi^2 = 82.7$, d.f. 3, $p<.0001$.

[78] JA: GB $\chi^2 = 56.8$, d.f. 3, $p<.0001$; IND: GB $\chi^2 = 97.5$, d.f. 3, $p<.0001$; IRE: GB $\chi^2 = 5.36$, d.f. 3, $p<.147$ (not significant).

uneven distribution of modal obligation/necessity can be related to regional discourse practices in the outer circle varieties. On the other hand, the present data seems to support Palmer's (1990 [1979]: 114) assertion that *have to* is more formal.[79]

Similarly to the private dialogues, *have to* is here the most preferred marker across all varieties, however, it found in different proportions. ICE-India displays a lower rate of this marker as compared to the face-to-face and telephone conversations. By contrast, *must* is slightly more often used here than in informal contexts with 114.3 occurrences per 100,000 words. This is however less surprising as this marker is associated with formal contexts of use, in general. The corpus-based observations seem to display a recurring feature in IndE which, impressionistically, appears to be a feature specific for this variety. Likewise, ICE-JA is placed second in this ordering even if it displays a striking preference for semi-modal *have to*. It is interesting to see that ICE-Ireland has surpassed both ICE-GB and ICE-India in the use of this semi-modal showing very low uses of *must* and even lower of *(have) got to* and *need to*. Nevertheless, it shows many similarities with ICE-GB.

As an interim conclusion, the data from ICE-JA, ICE-India and ICE-Ireland reveal stable layering within strong obligation and necessity. While the frequency of modal *must* is quite constant, the frequency of *have to* increases slightly with decreasing level of formality. An exception is semi-modal *need to* in ICE-JA which shows similar distribution with *must* across both the private and the public dialogues. Although the examination of the overall patterns of distribution with these forms has yielded significant differences both in the formal and the informal contexts of spoken interaction, it cannot be concluded that the four modal markers occur equally often in each text category. Such a configuration necessarily requires a separate analysis which is based on previous claims that suggest that the distribution of modals varies according to text type (Biber et al. 1999; Leech 2003; Mair & Leech 2006; Smith 2003).

The form-based distribution according to text type as listed in the tables 5.2, 5.3, 5.4 and 5.5 lends support to the observation that the correlations between the four modality markers and type of discourse are regionally specific. Perhaps more than in informal interaction the distribution of forms in the public dialogues points towards

[79] Note, however, that this assertion is based on the differences between *have to* and *(have) got to*.

specific regional preferences in the outer circle varieties, and which are not found in BrE.

As for informal contexts, the four modal items are biased towards frequent use in face-to-face dialogues. This is justified also by the smaller number of telephone conversations included in the corpus. If modal *must* (see Table 5.2) is overwhelmingly preferred in face-to-face conversations in all the four varieties with ICE-India at the forefront followed by ICE-JA, ICE-Ireland and ICE-GB, a similar pattern is found for telephone conversations, although in reversed ordering in the case of the latter two datasets:

		MUST			
		ICE-JA	**ICE-India**	**ICE-Ireland**	**ICE-GB**
Private	*Face-to-face*	113	177	111	86
	Phone-calls	20	32	7	10
Public	*Class lessons*	21	36	7	14
	Broadcast discussions	14	47	21	23
	Broadcast interviews	19	50	17	11
	Parliamentary debates	24	38	25	21
	Cross-examination	9	3	4	12
	Business transactions	15	9	8	6

Table 5.2. Distribution of modal *must* across private and public text dialogues

The higher rate of occurrences found in the 'phone calls' from the Indian dataset seems to suggest that speakers of Indian English are more confident to employ authoritative stance in a distant context of interaction. While according to Myhill (1995: 157) this item is typically employed in hierarchical social relationships, it is the functional analysis which will show whether it is the subjective evaluation (epistemic meaning) or rather strong obligation (root meaning) which prevails in our dataset (see Chapter 6). Likewise, the small number of texts included in this text category is likely to yield unreliable conclusions about the uses of obligation/necessity.

A comparison with semi-modal *have to* as occurring in the same text category shows that its frequency is approximately twice as high as that of the modal variant in ICE-JA and in ICE-India. This regularity in distribution in the two outer circle varieties

lends further support to the complex behaviour of the two markers *must* and *have to*. Apart from 'phone calls', 'parliamentary debates' represent another text category where the corpora display even distribution between *must* and *have to* across all the four varieties, as exemplified in (47) and (48). Unlike modal *must* in (47), semi-modal *have to* in (48) fulfils a performative function, which, as a recent development, seems to be a more common feature in ICE-JA:

(47) And I think <,> on a <,> on a bilateral basis particularly <,> we **must** make these friends feel and rewrite <,> that friendship in India <,> has always been very profitable <,> and will continue to be profitable for eachone of them <,> (ICE-India.S1B-054.txt)

(48) I **have to** thank my colleague Minister of State the Permanent Secretary and staff our advisors Ambassador King and others my personal staff and my secretaries (ICE-JA.S1B-053.txt)

Aside from informal contexts of interaction, ICE-India excels in uses of *must* in at least four other text categories pertaining to the section of public dialogues: 'class lessons' (36 occurrences), 'broadcast discussions' (47 occurrences), 'broadcast interviews' (50 occurrences) and 'parliamentary debates' (38 occurrences). Looking at the figures in Table 5.2 Jamaican speakers display a slightly lower but still similar number of occurrences.

As opposed to this trend, the class lessons from ICE-Ireland report a very low distribution of this marker, which is similar with the distribution of *(have) got to* and *need to* and in stark contrast with the high discourse frequency of *have to* (see Table 5.3). A further noteworthy point is the high incidence of modal *must* in the category of 'business transactions' in ICE-JA. Considering that interlocutors involved in such contexts usually aim at reaching an agreement that serves the interests of both parties, the choice for this marker of social normative and authoritative stance seems to be inappropriate in the context of highly specialized discourse. The more interesting question is whether the use of this marker in this text category from ICE-JA as compared to the other regional datasets shows signs of specialization of meaning.

		HAVE TO			
		ICE-JA	**ICE-India**	**ICE-Ireland**	**ICE-GB**
Private	*Face-to-face*	528	530	364	265
	Phone-calls	47	49	54	28
Public	*Class lessons*	89	79	83	33
	Broadcast discussions	83	87	91	65
	Broadcast interviews	54	29	31	37
	Parliamentary debates	31	38	25	16
	Cross-examinations	29	12	42	19
	Business transactions	83	43	41	52

Table 5.3. Distribution of semi-modal *have to* across private and public dialogues.

Moreover, the relatively even distribution of both *must* and *have to* in 'parliamentary debates' across the four corpora is indicative for a possible functional equivalence between the two modal items. Taking into account the argumentative structure of the genre itself, the modal would seem the natural choice at the expense of neutral and objective *have to*. The most remarkable contrast is found in ICE-India and ICE-Ireland where *must* and *have to* amount to 38 and 25 occurrences, as exemplified in (49) and (50):

(49) <#> But I 'd have the House to note that the Government 's position is <,> it **must** be based on a comprehensive political settlement commanding widespread support (ICE-Ireland. S1B-051.txt)

(50) So now the time has come <,> we are not to <,> think twice that <,> we **have to** give this top priorities for the export of the agricultural commodity <,> (ICE-India.S1B-053.txt)

While JamE, IndE and IrE show an abundance of occurrences with *have to* in 'class lessons' and 'broadcast discussions' a lower frequency of use in these categories from ICE-GB might possibly denote different discourse strategies. Interestingly, the text category 'cross-examinations', in which modal *must* would be expected to be frequently employed, displays a lower range as compared to semi-modal *have to*. In view of the general purpose of cross-examinations – i.e. to question somebody in detail and to extract as much information about a legal issue – authoritative stance is expected to be

employed more often with *must* as a logical choice to express obligation in institutionalized contexts of this type. This does not seem to be the case in our four datasets as it is also illustrated in (51) taken from ICE-Ireland where semi-modal *have to* mitigates the effect of the authority in a legal court:

(51) if I 'm correct you were sitting at a station and this car came along at twenty miles an hour and crashed into your rear <#> You can't say that <#> You **have to** give the reference not me <#> And he's to ask the questions in a particular way <#> He can't lead you <#> So that's why he's asking you in this rather official way (ICE-Ireland.S1B-061.txt)

On the other hand, the semi-modal appears to be the natural choice if at the centre of debate are court procedures or instructions (see Depraetere & Verhulst 2008). In this sense, the use of this semi-modal is part of a strategy to maintain objectivity which accommodates the public context of such a text category. Likewise, one should not overlook that this semi-modal also reports past necessity. In fact, semi-modal *have to* is overall the favoured option for this text category with most occurrences (42) to be found in ICE-Ireland followed by ICE-JA (29). Even so, it is the modal *must* which is most favoured in this text category in ICE-GB (12 occurrences) as compared with the other datasets.

		(HAVE) GOT TO			
		ICE-JA	**ICE-India**	**ICE-Ireland**	**ICE-GB**
Private	*Face-to-face*	8	9	13	113
	Phone-calls	4	0	1	18
Public	*Class lessons*	2	0	8	19
	Broadcast discussions	3	3	20	10
	Broadcast interviews	0	0	11	15
	Parliamentary debates	0	3	2	0
	Cross-examinations	0	0	0	3
	Business transactions	1	0	1	9

Table 5.4. Distribution of semi-modal *(have) got to* across private and public dialogues

As already discussed in the comparison of the overall frequency patterns, semi-modal *(have) got to* (see Table 5.4) and *need to* (see Table 5.5) are underrepresented in the

present database. It is only ICE-Ireland which shows consistent use of *(have) got to* in at least two text categories, namely 'broadcast discussions' as in (52) and 'broadcast interviews' as in (53):

(52) <#> Maybe on either side people are realising you **'ve got to** to resolve this problem one way or another <,> and an attempt is being made now through the the the Downing Street Declaration the Peace Process to uh resolve it and certain uh home truths have been accepted on on all sides or on some of the sides anyway (ICE-Ireland.S1B-021.txt)

(53) <#> I was yeah <#> Uhm I think I was uh definitely apprehensive about going over but I just said to myself <,> I **'ve got to** give it a try (ICE-Ireland.S1B-047.txt)

Judging from the examples above, strong subjective/emotional obligation expressed with this semi-modal accommodates well such text categories particularly in the Irish context.

		NEED TO			
		ICE-JA	**ICE-India**	**ICE-Ireland**	**ICE-GB**
Private	*Face-to-face*	121	16	11	46
	Phone-calls	12	0	3	8
Public	*Class lessons*	49	3	8	25
	Broadcast discussions	14	16	8	10
	Broadcast interviews	11	1	2	8
	Parliamentary debates	4	1	5	16
	Cross-examinations	1	0	1	5
	Business transactions	38	1	7	12

Table 5.5. Distribution of semi-modal *need to* across public and private dialogues

As regards semi-modal *need to* (see Table 5.5), apart from its preference in informal contexts in ICE-JA, it is most often employed in 'class lessons' (54) and 'business transactions' (55):

(54) <#>These are errors that proof reading very often would help to correct right so you **need to** leave time at the end but I think you **need to** go back over this first part <}><->of the of the</-> uhm<,> <=>of the</=></}> outline (ICE-JA.S1B-005.txt)

(55) <#>We **need to** be looking at the schedule of benefits that are in relation to those jobs at present because there might be other things (ICE-JA.S1B-073.txt)

ICE-JA contrasts in this sense with a similar pattern in ICE-GB, however, the latter dataset showing preference for this item mainly in 'class lessons' and 'parliamentary debates'. The specific distribution of *need to* in ICE-JA is worth mentioning, as it seems to enter direct competition with both the semi-modal *have to* (see Table 5.2) and the modal *must* (see Table 5.1).

To sum up, the examination of distribution patterns according to text categories supports the assumption that the strongest alternation in the three New Englishes – JamE, IndE and IrE – is found between the modal *must* and the semi-modal *have to* both across informal and formal text-categories.

5.2.2 Morphosyntactic and stylistic characteristics in New Englishes

5.2.2.1 Syntactic motivation: have to *and* need to

According to Leech et al. (cf. 2009: 80), there is a link between the recent decrease in the use of modals, also referred to as "paradigmatic atrophy", and their lack in morphological contrast in person and number, which makes modals appear as "anomalous" or "defective" verbs. It seems logical to attribute the high incidence of semi-modal *have to* and *need to*, particularly in ICE-JA, to such morphosyntactic features.

It was mentioned earlier in 4.6.2 that semi-modal *have to* has the advantage of occurring in syntactic environments in which modal *must* is not available whereas *(have) got to* is hardly ever found (see Hundt 1998: 55; Johansson & Oksefjell 1996; Krug 2000: 103). Similarly, semi-modal *need to* has grown to replace modal *need(n't)* in affirmative and declarative utterances. In addition, the rapid spread of semi-modal in present-day spoken and written English is often attributed to their morphosyntactic flexibility. Following Myhill (1995: 166) and Smith (2003: 254), four structural environments differentiate *have to* from *must*:

(a) non-finite forms (e.g. *having to, to have to*)

(b) following a modal (e.g. *will have to*)

(c) negated forms expressing absence of requirement (e.g. *don't have to*)

(d) past tense (e.g. *had to*)

The same would be true for semi-modal *need to* which occurs both in assertive and non-assertive contexts. An assessment of such claims in the present data considers the extent of these characteristics across JamE, IndE and IrE.

Table 5.6 and 5.7 summarise these distinctions for semi-modal *have to* and *need to* in the spoken dialogues from the present dataset. Note, that although there is consistent use of syntactically motivated uses with *have to* (see Table 5.6), the non-syntactic uses prevail overall:

HAVE TO	**ICE-JA**	**ICE-India**	**ICE-Ireland**	**ICE-GB**	**SBC**
Syntactically motivated	239 (66.3)	184 (51.1)	300 (83.3)	243 (67.5)	248 (99.5)
Non-syntactically motivated	714 (198.3)	680 (188.8)	434 (120.5)	286 (79.4)	304 (122)

Table 5.6. Distribution of syntactic vs. non-syntactic uses of semi-modal *have to* in spoken dialogues from ICE-JA, ICE-India, ICE-Ireland, ICE-GB and SBC. The values in parentheses indicate normalized frequencies per 100,000 words.

Interestingly, the largest proportion of non-syntactic *have to* is found in ICE-JA, ICE-India and ICE-Ireland, even though in the latter dataset the gap between the two types of uses is smaller than in the former two. Contrastively, the rate between syntactic and non-syntactic uses in ICE-GB and SBC is narrower than in the other varieties and which would confirm Krug's (2000: 95) observation regarding the rise of such uses both in present-day BrE and AmE. Moreover, although the first obligation readings with *have to* emerged out of present tense forms (Krug 2000: 96) it is difficult to account for either retention or innovation in the use of this semi-modal in the two outer circle varieties. Nevertheless, the high incidence of such uses in the spoken dialogues from ICE-JA and ICE-India lends further support to the functional layering within strong obligation/necessity particularly in outer circle varieties.

Although the alternation between the non-syntactic uses with *have to* and *must* in ICE-GB and SBC does not appear to be very pronounced, their distribution (see Appendix 3.a) shows that occurrences with the semi-modal outnumber by far those with the modal. Statistical Chi-square testing using Yates' correction (Oakes 1998: 25) for

binomial distributions confirms very high significance at $p<.0001$ only for ICE-JA as correlated with ICE-GB, and for ICE-India with ICE-GB and SBC.[80] In statistical terms, there are marked differences between the syntactic and non-syntactic uses of *have to* in JamE as correlated with BrE, but not with AmE as $p=.4463$ (not significant). In contrast to JamE, the distribution of syntactic and non-syntactic uses in IndE differs markedly in correlation with both BrE and AmE. Moreover, there are no statistically marked differences in the distribution found in IrE as correlated with BrE and AmE. That is, in spite of the overall larger number of *have to* in ICE-Ireland as compared to ICE-GB and SBC, the speakers of these three varieties do not vary much between the two types of use.

As regards *need to*, the contrast between syntactic and non-syntactic uses is even more pronounced (see Table 5.7) across the five varieties:

NEED TO	**ICE-JA**	**ICE-India**	**ICE-Ireland**	**ICE-GB**	**SBC**
Syntactically motivated	11 (3)	10 (2.7)	33 (9.1)	29 (8)	19 (7.6)
Non-syntactically motivated	239 (66.3)	28 (7.7)	74 (20.5)	93 (25.8)	103 (41.3)

Table 5.7. Distribution of syntactic vs. non-syntactic uses of semi-modal *need to* in spoken dialogues from ICE-JA, ICE-India, ICE-Ireland and SBC. The values in parentheses indicate normalized frequencies per 100,000 words.

Despite these fluctuations, statistical significance using Yates' correction emerges for ICE-JA when correlated with ICE-GB at $p<.0001$ and with SBC at $p<.0004$, as well as for ICE-Ireland, but only when correlated with SBC at $p=.00095$.[81] It is IndE which shows no statistically verifiable differences in the patterns of use with *need to* as correlated with the two native varieties.

Overall, the examination on the basis of syntactic factors of *have to* and *need to* shows that non-syntactically motivated uses can be considered as a source of variation

[80] ICE-JA: ICE-GB $X^2 = 66.48$, d.f. 1, $p<.0001$; ICE-India: ICE-GB $X^2 = 92.55$, d.f. 1, $p<.0001$; ICE-Ireland: ICE-GB $X^2 = 3.01$, d.f. 1, $p<..0828$ (not significant); ICE-JA: SBC $X^2 = 0.58$, d.f. 1, $p<.4463$ (not significant); ICE-India: SBC $X^2 = 87.61$, d.f. 1, $p<.0001$; ICE-Ireland: SBC $X^2 = 1.96$, d.f. 1, $p<.1615$ (not significant).

[81] ICE-JA: ICE-GB $X^2 = 30.07$, d.f. 1, $p<.0001$; ICE-India: ICE-GB $X^2 = 0.01$, d.f. 1, $p=.9203$ (not significant); ICE-Ireland: ICE-GB $X^2 = 1.11$, d.f. 1, $p<.2921$ (not significant); ICE-JA: SBC $X^2 = 13.34$, d.f. 1, $p<.0004$; ICE-India: SBC $X^2 = 1.59$, d.f. 1, $p<.2073$ (not significant); ICE-Ireland: SBC $X^2 = 6.73$, d.f. 1, $p=.00095$.

in the spoken dialogues in ICE-JA, ICE-India and ICE-Ireland. To sum up, this suggests that the use of strong obligation/necessity is quite circumscribed in the three focus varieties and resembles the case of non-standard varieties of BrE described in Tagliamonte (2004: 40).

5.2.2.2 Not negation *vs.* do negation

Another feature of the layering in the system involves negated forms of modal expressions: i.e. absence of requirement (*do*-support) as opposed to negation of the proposition (*not*); or between *external* and *internal* negation (Huddleston & Pullum 2002: 204). Table 5.7 summarizes the findings of negative polarity in the spoken data:

(Semi-)modal	ICE-JA	ICE-India	ICE-Ireland	ICE-GB	SBC
must not/mustn't	8 (2.2)	5 (1.3)	6 (1.6)	14 (3.8)	0
need not/needn't	2 (0.5)	15 (4.1)	8 (2.2)	1 (0.2)	0
*NOT *have to*	31 (8.6)	24 (6.6)	30 (8.3)	27 (7.5)	60 (24)
have not got to/ haven't got to	0	0	0	2 (0.5)	0
*NOT *need to*	8 (2.2)	3 (0.8)	2 (0.5)	7 (1.9)	9 (3.6)

Table 5.8. Distribution of negated forms in spoken dialogues from ICE-JA, ICE-India, ICE-Ireland, ICE-GB and SBC. The values in parentheses indicate the normalized frequencies per 100,000 words.
*Stands for *do not, does not, did not, don't, doesn't, didn't*.

Due to the overall low frequency counts of negated *must* and *need*, negative contraction (*-n't*) is not treated separately in the table. Likewise, negation with modal *need* expressing absence of necessity is being considered here mainly because it may function as suppletive for negated *must* (Nokkonen 2006: 65), as well as to emphasize its rare use in the database altogether, as compared to *do*-support with the semi-modals *need to* and *have to*. The scarcity of such forms is, then, in accordance with the assumed "paradigmatic atrophy" which characterizes modal verbs in general (Leech et al. 2009: 81-82). Despite this irregularity, the distribution of negated forms with the marginal modal *need* occurring in ICE-India and ICE-Ireland is noteworthy. Only one occurrence of *needn't* is reported in ICE-GB, whereas in SBC this form does not occur at all. The

IrE private dialogues provide three occurrences of which (56) is an example where the speaker imposes a certain type of behaviour using the less authoritative form of negated *need*, and which is typically found in spoken language (Nokkonen 2006: 65):

(56) She was saying something about <,> well you know at least at least it 's so cool like you know when you 're driving round Dublin <,> doing your driving test you know <,> you **need***n't* ever go any faster than thirty-five miles an hour (ICE-Ireland. S1A-064.txt)

Decrease in the use of negated *need* concords with previous findings such as reported in Krug (2000: 202-203) in data from BNC. The same observation can be extended for *must not/mustn't*, which, although low in frequency, seems to be more common in BrE (3.8 occurrences per 100,000 words) and JamE (2.2 occurrences per 100,000 words). This is a striking finding taking into account the overall preference of this modal, in general, in the IndE dialogues. The lack of this type of negation in SBC is in accordance with the recent trends, namely that *must* is on its way out of the modality system in AmE. While most of the contractions with negated *must* are more common in informal English, this form is found also in formal contexts of interaction as in (57) extracted from a 'class lesson' from ICE-Ireland, in which the lecturer gives instructions to students for a specific assignment in a sociolinguistics project,[82] and in (58) extracted from a broadcast discussion from ICE-IND:

(57) <#> Uh and there 's a lot of choice that you can use within that <#> The main thing is it 's naturalistic <#> It **must***n't* be faked <,> it mustn't be acted <,> it mustn't be off TV <,> it mustn't be off radio <#> Uhm and believe me if it 's any of those things I 'll detect it immediately (ICE-Ireland.S1B-001.txt)

(58) Uhm the fact matter because <,> you can't uh wish away I mean the Soviet Union <,> is not irrelevent I mean <w> it's </w> <,> one **must***n't* obtained that after the break up of the Soviet Union <,> Russia has become irrelevent (ICE-India.S1B-039.txt)

Interestingly, example (58) illustrates a feature already commented in Westney (1995: 138-139): whether there is always semantic correspondence between *must not*/*musn't* and positive polarity with *must*, given that this is a default feature of the class of the English modal verbs. According to Westney (1995: 139), the existence of alternative

[82] Although there are three occurrences of negation with contracted *must* in example (47), it was counted as one token, whereas the other two occurrences were coded as repetitions within the same speech unit.

constructions (e.g. *have to*, *have got to*) expressing negative polarity of modal necessity in present-day English lends *must not*/*musn't* to be employed for positive polarity of other variant forms than *must*. With this observation in mind, *mustn't* in (58) appears to correspond to positive contexts of neutral *have to*, as the negation expressed here is less directive and authoritative, and seems to denote rather non-requirement of necessity. Such use might be evidence for a further stage in the development of this modal. However, the low number of such occurrences does not allow for any strong claim of a generalized pattern.

Although this section addresses form rather than function, the epistemic use with *mustn't* in excerpt (59) taken from a face-to-face conversation is a typical phenomenon in IrE and a range of other standard and non-standard varieties (see Corrigan 2000; Hickey 2004a; 2007):

(59) <S1A-009$B> Well Gerry didn't know that Esther Dunne was any relative
<S1A-009$A> <#> Oh did he not
<S1A-009$B> <#> No <#> He never heard his mother speak of her <#> He didn't know her <,> or anything about her <#> And his mother was Geraldine McSwiggan
<S1A-009$A> <#> Oh I know that
<S1A-009$B> <#> Well Esther Dunne **must***n't* have been there at it then <#> I said our Rosie says she 's related to him
<S1A-009$A> <#> Oh yes <#> In some way Caroline yes (ICE-Ireland.S1A-009.txt)

Apart from the inferential interpretation, further features support the epistemic reading of the proposition, such as the perfect aspect which reports past reference as well as the temporal adverb *then* (see also section 6.4 on epistemic necessity).

Conversely, consider example (60) from a broadcast discussion from ICE-JA where one of the speakers uses modal negation with *need* denoting absence of necessity. In the new conversational turn another speaker engages in the discussion using the same expression followed by a marked infinitival complement which is incompatible with the negated structure expressed with modal *need*:

(60) <$C><#> So you're saying as long as it's for your husband then he **need not** come<$E><#>He **need not** *to* come because when the town got <unclear>words</unclear> that my husband should come with me <.>dow</.> go down there with me <#>When we go they said he *does***n't have to** come <?>yeah anyone</?> could just come and do it because all the papers

<unclear>word</unclear> and originals are all there so when I gonna just look for that file I'd go <,>What you have to do (ICE-JA.S1B-036.txt)

While this, too, is an isolated example indicating a false start, as the speech sequence includes also unclear passages, this instance is quite ambiguous and therefore was not included in the counts from table 5.3. However, some noteworthy observations call for discussion. Apart from the hesitations and the interruptions, it is also the choice for the contracted future marker *gonna* which confers the speech context here an informal character. Interestingly, at a later point in the flow of uninterrupted speech the same speaker expresses absence of negation this time using the equivalent structure NOT *have to* followed by the same dynamic verb 'come', which confirms the existence of a highly layered system.

At the opposite end, negated *need to* denoting external requirement of non-performance of an action (Taeymans 2004: 105) is quite infrequent. The lowest distribution is found in ICE-India, which is in stark contrast with the distribution of its modal counterpart as well as with the semi-modal *have to*. Another nice example of alternation between negated *need to* and *have to* is found below in (61):

(61) <#><[>That's true but</[></{> when the course starts out <}><->you might not have</-> <=>you might be</=></}> understanding the course so you feel you *do***n't need to** go to the lecturer but just for your sake only you need to go to him and say listen my name is X and I'm in your class and thereafter you just find small items just to go to him and go to <#>You *do***n't have to** go every week<{><[><,></[> ICE-JA.S1A-058.txt)

Some noteworthy observations emerge from this example. In the first speech unit the choice for negated *need to* seems to be internally motivated as signalled by the perception verb *feel*, whereas in the second instance negated *have to* denotes non-fulfilment of a habitual action, and is emphasised with the adverb *every*. It follows from this that, at least in this speech unit, the speaker is aware of the distinct semantic load that each of the two semi-modals carry.

Overall, negation with semi-modal *have to* is the most common form to express absence of requirement across all the four varieties. As expected, AmE displays the highest frequency rate whereas the other three varieties show a balanced pattern with

minor fluctuations in distribution. Moreover, the lower frequency of occurrence of negated *have to* in ICE-India shows more similarities with the British pattern.

Finally, two negated forms of *(have) got to* were found to be used only in ICE-GB, as in example (62), which can be compared with the small number of this structure found in the BNC (Krug 2000: 108):

(62) That's all very well isn't it when you've got <,> cases where you **have***n't* **got to** worry about (ICE-GB.S1B-010.txt)

Utterance (62) extracted from the text category 'class lessons' is striking not only for the rare use of the negated form in present-day English, but also for the abundance of contracted auxiliaries co-occurring in one speech unit, which are more common for informal settings. This is suggestive for the semiformal character of the text category 'class lessons' which I suggested in section 4.7.2.

To sum up, the examination of negation patterns shows that the semi-modal *have to* is, indeed, moving to become the most common form of obligation and necessity in JamE, IndE and IrE.

5.2.2.3 Contractions

An even more differentiated illustration of the synchronic behaviour of the semi-modals of necessity in the New Englishes can be retrieved from the assessment of stylistic features which involve contracted forms, e.g. *'ve to/'ve got to/'s got to/gotta*. According to Krug (1998b; 2000), *string frequency* is a central cognitive factor motivating the process of univerbation and ongoing language change of both the semi-modal *have to* and *(have) got to*. Moreover, in spite of the non-standard origin of *(have) got to*, the emergence of this construction and its phonologically reduced variants as informal and s></urls></record></Cite></EndNote> (e.girectly related to increase in discourse-frequency (Krug 2000: 62-64). According to this notion, the highest the incidence in discourse frequency of an element the more probable it is that it undergoes univerbation.[83] This process entails coalescence and cliticization to adjacent linguistic

[83] The notion of *string frequency* goes back to Zipf's (1929 ff.) Principle of Least Effort, which assumes a correlation between frequency of occurrence and morphological simplicity, and between frequency and the age of words (Krug 1998b: 286 f.).

forms, for example, of semi-modal *have to* to *'ve to* and *have got to* to *'ve got to/'s got to* as well as morphological reduction to *gotta*.

More remarkable for the present discussion is that contractions are an undeniable source of ongoing change in present-day English as related to formality levels. It is style which, apart from frequency constraints, is a conditioning factor in the emergence of such forms both in written and spoken language (see Diller 1999; Kjellmer 1998; Krug 1998b; 2000; Mair 2009b). Thus, a higher rate of contractions or phonologically reduced forms would involve increased stylistic variability, which in this case would mean a tendency towards the informal style.

So far, the data indicates stable use of *have to* as opposed to disproportionate uses of *(have) got to*, as the most informal alternative to express obligation in ICE-JA, ICE-India and ICE-Ireland. As shown in Table 5.8, the present dataset builds on a twofold contradiction:

Semi-modals	**ICE JA**	**ICE Ind**	**ICE Ireland**	**ICE GB**	**SBC**
've to	0	0	51 (14.6)	1 (0.2)	0
've got to/'s got to	2 (0.5)	4 (1.1)	25 (6.9)	136 (37.7)	7 (2.8)
gotta	6 (1.6)	0	1 (0.2)	0	101 (40.5)

Table 5.9. Frequency of contracted variants for *have to* and *have got to* in the spoken sections of the five corpora. The values in parenthesis indicate the normalized frequencies per 100,000 words.

Although contracted forms are more common for the conversational style and although *have to* is most frequent in JamE and IndE, it is particularly in these varieties where encliticization of *'ve to* is missing. By contrast, the dialogues in ICE-Ireland are replete with such features displaying 51 raw modal instances of *'ve to*, which represent 9.6% of the total amount of entries of semi-modal *have to* in ICE-Ireland as well as 25 raw instances of *'ve got to/'s got to*, which represent 43.8% of the total amount of this construction (see Appendix 3.a). More surprising is the fact that ICE-GB and SBC display only one entry, and no entry of *'ve to*. Even though the majority of such contractions in the Irish dataset are found in the private dialogues (N=43; 84.3% of the total amount of contracted *have to* in ICE-Ireland) as illustrated in (63), formal contexts of speech prove to be also a potential source for variability, as in (64):

(63) <#> There are twelve lectures I **'ve to** get through <,> each of them consisting of four pages right <,> well two pages back and front which is four pages <,> the whole thing filled out nothing like no gaps or lines or anything like that (ICE-Ireland. S1A-075.txt)

(64) <#> What what I 've been doing down in Henry Flood has been <,> I 've asked Elaine and Jim to do the supervised training but if I have a spare half an hour or something or if I 'm waiting to see somebody and I **'ve to** wait for them to get sitting up in bed or something like that and there 's a nurse I 'll say do you want to go and do a supervised training now (ICE-Ireland.S1B-078.txt)

Overall, the spread of the form *'ve to* in the Irish formal contexts of speech is biased towards uses in 'class lessons' (2 occurrences), 'broadcast discussions' (1 occurrence) and in 'business transactions' (3 occurrences).

Despite the overall low frequency of the semi-modal *(have) got to* in the three varieties under study, this expression is found in the database as the variant *got to*/*'ve got to*/*'s got to* and *gotta*. From this perspective, the Jamaican dialogues represent an untypical case. Contrary to expectations, ICE-JA shows full range of synchronic layering of semi-modal *(have) got to*, as all its variant forms are present, e.g. *'ve got to,'s got to, got to* and even *gotta*, which potentially suggests movement in the system. In fact, the occurrences of *gotta* represent 35.2% of the total amount of semi-modal *(have) got to* in the Jamaican dialogues (see Appendix 3a), and which are spread equally in private (65) and public conversations (66):

(65) <#>I walk a lot in the mornings right <?>I jog</?> and <}><->I in</-> <=>in</=></}> July and August I **gotta** get home back by nine o'clock (ICE-JA.S1A-005.txt)

(66) <#>Him seem to know everything about all that now<{><[><O>laughs</O></[><$C><#><[><unclear>words</unclear><O>laughs</O></[></{> <#>Oh boy <#>We **gotta** take a break but coming up is James Jimmy solo Howard is gonna be joining us (ICE-JA.S1B-036.txt)

However, since the raw figures are very small, such an assumption is based on speculations, and further research is necessary to gain a reliable representation of such variability.

One aspect on which BrE and AmE seem to converge as concerns the present dataset is the high discourse frequency with the informal variant forms of *(have) got to*. However, the data also shows that the two reference varieties contrast significantly in

the preference to use '*ve got to* in BrE and *gotta* in AmE, thus, giving credit to Kurg (2000: 72) with regard to the development of the grammatical inventory in present-day English.[84]

So far, the data listed in the table above (see Table 5.8) is insufficient to unravel stylistic variation in the outer circle varieties. Moreover, judging from the numbers shown above, it is tempting to label the use of contracted *'ve to* as a local phenomenon in IrE. However, in the absence of evidence in the literature with regard to such uses in IrE, and in particular in non-standard Irish varieties, such an interpretation calls for caution.[85] Although the compilation of the national components of ICE required the adoption of the same transcription and markup systems, a further issue involves accuracy in the transcription of the spoken conversations.

Furthermore, while the distribution analysis of contracted *(have) got to* in the Jamaican dialogues can be accounted for by Rohdenburg's (1996: 173) *complexity principle* as invoked in Krug (2000: 110-111),[86] the lack of enclitization with *have to* in JamE, IndE, BrE and AmE does not tally with the notion of *string frequency*. Even if *have to* is the most preferred marker of obligation and necessity, stylistic variation as correlated with discourse frequency does not seem to account for ongoing development in the two outer circle varieties, JamE and IndE.

5.3 Distribution of forms in written press texts

Although the main focus in the present study is on spoken dialogues, a comparison with the written medium potentially strengthens the interpretation of the common synchronic trends shared across the outer and inner circle. Thus, this section discusses the distribution of forms of *must*, *have to*, *(have) got to* and *need to* in press texts as representing formal writing (for details on the selection of text categories, see section 4.3.2). Particularly BrE and AmE journalistic prose are likely to incorporate more easily

[84] Central to Krug's (2000: 72) argumentation is Gabelentz's (1891: 251) spiral model of language change.
[85] I thank Karen Corrigan (personal communication) for drawing my attention on this caveat.
[86] The *complexity principle* states that: "In cases of syntactic variation the more explicit option is generally more formal than its less explicit counterpart" (Rohdenburg 1996: 173). According to Krug (2000: 110-111), due to its phonological structure modalized *gotta* emerges as the most explicit syntactic option as opposed to *have to* accommodating thus the complexity principle.

linguistic innovation such as increased use of colloquial language, which places this type of writing among the 'agile' genres (Hundt & Mair 1999). This trend is often described in terms of the narrowing gap between spoken and written language as a shift in stylistic preferences. The present working hypothesis is that outer circle varieties are expected to apply the 'educated' norm more strictly as the target in written texts than inner circle varieties. It is therefore instructive to assess whether these modality markers vary regionally to the same extent in the written as in the spoken texts.

As it becomes apparent from Figure 5.4, the range of variation in the press texts differs from that in the spoken texts. Note also that the press texts in ICE and Frown are smaller in size than the spoken dialogues analysed in the previous section.

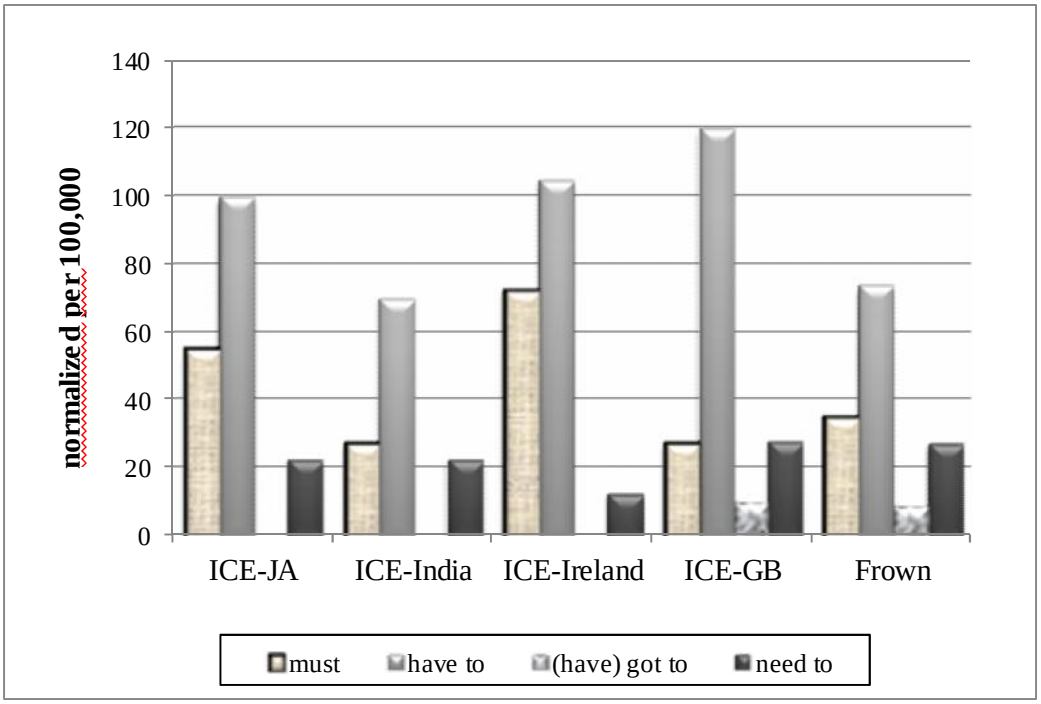

Figure 5.4. Distribution of strong obligation/necessity per 100,000 words in written press texts

Figure 5.4 displays the frequency of forms as normalized per 100,000 words, whereas a detailed overview of all raw frequency counts in the press texts is provided in Appendix 3b. Unlike the case of spoken texts, I do not report significance of Chi-square values for press texts, as many cells contain values smaller than 5 and even zero, and do not fulfil the minimum requirement for such tests. This is particularly the case of the semi-modal *(have) got to* in ICE-JA, ICE-India and ICE-Ireland.

The most striking difference in this representation is found with ICE-Ireland, where modal *must* displays the highest incidence (72.5 per 1000,000 words) across all the four datasets. In fact, the frequency of the modal *must* in ICE-Ireland outnumbers semi-modal *have to* in ICE-India, where this marker holds the highest position in ranking. Thus, modal *must* is losing ground in ICE-GB, Frown and ICE-India – and where its use would be expected to occur more often – as compared to the high frequency found in ICE-JA and ICE-Ireland. This result is intriguing considering that the Indian spoken dialogues have shown robust use of modal *must* as compared with the other varieties, whereas the press texts seem to follow a different trend, such as the "colloquialization" of the written norm.

The frequencies of *have to* are remarkably stable, however displaying minor variation. Contrary to expectations ICE-GB (120 occurrences per 100,000 words) is on the forefront with regard to the preference of this semi-modal in the press texts. Very close to this ranking are ICE-Ireland (105 occurrences per 100,000 words) and ICE-JA (100 occurrences per 100,000 words). Thus, taking discourse frequency of *have to* as evidence for the stylistic shift in written language, at least the Irish and Jamaican press texts appear to follow the more colloquial, direct speech style among the three New Englishes. Furthermore, from Figure 5.4 it emerges that BrE is more advanced as regards the use of strong obligation/necessity in journalistic prose. In support of such evidence, it will be useful to recall the overall trend of this item observed in journalistic prose in the Brown corpora (Krug 2000: 84). Considering Krug's findings, the present dataset would indicate that BrE is more advanced than AmE in the use of this semi-modal.

On the other hand, semi-modal *need to* shows small frequency counts whereas *(have) got to* does not occur at all in ICE-JA and ICE-India. As for ICE-GB and Frown, both display a similar pattern in the dispersion of *(have) got to*. Further interesting parallels can be found at least as regards the uses of *must*, *have to* and *need to*: the data from ICE-Ireland and ICE-JA show more similarities with ICE-GB, whereas data from ICE-India rather with Frown. As for *(have) got to*, there is clearly a difference to be noticed as compared to the spoken data. Again, it is ICE-GB and ICE-Ireland where this construction is reported, however with low frequency, namely four occurrences in the former and only one in the latter case. The inexistence of *(have) got to* in the outer

circle and IrE tentatively reinforces the conjecture of strong prescriptive ties followed in press texts in these communities.

In broad terms, the analysis of the press texts reveals, that unlike the spoken texts, the outer and inner circle seem to converge in regard to the written norm. However the steady increase of *have to*, for example in the British, Irish and Jamaican data can be considered as a sign of text specialisation as corroborated in previous findings from LOB and FLOB (Smith 2003: 252).

5.4 Summary

The quantitative analysis in this chapter has concentrated on the distribution of obligation and necessity across spoken dialogues and press texts in three different varieties – JamE, IndE and IrE – as compared to BrE and AmE. Therefore the function-based analysis was not included here. The purpose of this chapter has been to identify the major synchronic tendencies in the use of these modality markers in standard JamE, standard IndE and standard IrE and to establish the extent of variation on three interrelated dimensions: (a) medium-specific and stylistic; (b) regional and (c) diachronic. The spoken data was further analysed on its morphosyntactic and stylistic characteristics. Frequency of occurrence was considered a significant factor for identifying the grammatical status of strong obligation and necessity in different regional settings.

So far, the data has provided interesting patterns in the regional distribution of these markers. It emerges from this analysis that the spoken dialogues in the three focus varieties JamE, IndE and IrE are a potential source for variablity within modal obligation and necessity, whereas their use in press texts converges with the written norm. Moreover, the comparison of the ICE corpora shares a lot of similarities with data from other English varieties. Thanks to the numerous corpus linguistics studies undertaken on this subject there are some general tendencies which can be traced back in our data as well.

Although competition between modals and semi-modals in present-day English is a widespread phenomenon, it has proved difficult to correlate the synchronic

behaviour in the outer circle with diachronic treatments in BrE and AmE. On the whole, the data shows variation on all levels in the Jamaican, Indian, and Irish components of ICE, and which is more pronounced in speech than in the newspaper reports. JamE and IndE, clearly form a pattern especially as regards the distribution of the modal *must* and the semi-modal *have to* in spoken interaction. Likewise, they display a similar pattern in the under-representation of colloquial/informal *(have) got to*. Especially *have to* seems to gain ground in Jamaican and Indian speech whereas usage of modal *must* is robust, although significantly at a lower rate of frequency. On the other hand, in the inner circle varieties (IrE, BrE, and AmE) this marker seems to be used moderately, signalling a possible new direction of development.

As regards contractions, which in spontaneous interaction signal rather informal speech habits, both IrE and BrE are leading, whereas AmE shows a clear tendency towards using the most innovative variant *gotta*. However, as I argued in section 5.2.2.3, due to insufficient evidence it seems premature to conclude that IrE is more advanced than JamE and IndE in using contractions. Moreover, it is also interesting that British English exhibits contracted forms only for semi-modal (*have) got to* (‘*ve got to*/’*s got to*). The relatively high distribution of *must* in newspaper reports confirms the assumption with regard to retention of the dominant British written norm in Jamaican and Irish texts, and contrasts with a similar frequency of the newer layer *have to* in the Jamaican, Irish, and British texts. At this stage of the analysis, it is Indian English which appears to be more conservative.

Overall, despite common trends, the discrepancies in the distribution and frequency patterns found in JamE, IndE and IrE suggest that differential speeds of change are at work as compared with BrE and AmE.

6 Semantic and contextual analysis

6.1 Introduction

In the previous chapter, I presented an overview of the main trends in the distribution of strong obligation/necessity in JamE, IndE and IrE as compared to BrE and AmE. The main focus was on the overall frequency patterns and the way this information contributes to the understanding of the synchronic behaviour of obligation and necessity in the New Englishes. I also showed that the variable distribution in the spoken components of ICE-JA, ICE-India and ICE-Ireland points towards gradual dispersion of these items: with the increase in frequency of occurrence contrasting with a decrease in the formality level of the text categories. In addition, frequency of occurrence has revealed a significant statistical difference in the ordering of these items both in conversations and written press texts within the outer circle (as opposed to the inner circle varieties). The analysis reveals that in terms of overall frequencies most alternation occurs between the modal *must* and semi-modal *have to* across the three varieties under study.

Although the quantitative analysis has particularly emphasised the synchronic layering of forms in the New Englishes in terms of normalized frequencies, it cannot be considered conclusive enough to warrant ongoing change within the domain of obligation and necessity. To this end, in this chapter I will concentrate on semantic contrasts with *must*, *have to*, *(have) got to* and *need to*. The analysis aims to uncover whether the variable ordering of the overall frequencies as discussed in Chapter 5 is related to a difference in the qualitative range of modal functions with these markers.

There are two specific claims which will be tested. The first one refers to the generalization that one can infer ongoing change in modal necessity/obligation in JamE, IndE and IrE, as compared to BrE and AmE, by evaluating distribution patterns of semantic contrast in context. The other is more specific and refers to earlier statements in the literature on the modal – semi-modal alternation, namely that there is no clear distinction related to the subjective use of *must* and necessity, and *have to*, *(have) got to* or *need to* and objective necessity (see also section 1.2). Building on the fact that the various uses of these items in present-day English range from very specific to general meanings the other level of analysis consists of establishing in which way the present

data conveys subjective or objective root necessity as used in Coates (1983) and Depraetere & Verhulst (2008) (see section 6.3). From this perspective, semantic contrast serves as an essential tool to disambiguate various uses, and, ultimately, to uncover the observed asymmetries in frequency.

The semantic interpretation of the data follows the binary distinction of root vs. epistemic reading (see section 4.5.1). As in the previous chapter, semantic contrast will be assessed in light of previous findings from BrE and AmE as well as from other non-standard and non-native varieties of English. As shown in section 5.2.2.1, alternation in spoken interaction between modal *must* and semi-modal *have to* or *need to* is limited to non-syntactic environments. Therefore, a true comparison between these forms will be approached by only analysing semantically motivated morphemes (Myhill 1995: 166).

In view of the dominant use of the semi-modal *have to* across all the five varieties – in particular in data from JamE, IndE and IrE – the present examination might have further consequences on the generally accepted realization of the root vs. epistemic contrast with this marker in New Englishes (cf. Westney 1995: 101).

6.2 Semantic contrast in the New Englishes: root vs. epistemic uses

6.2.1 Spoken dialogues

The present section focuses on the description of distribution patterns of root and epistemic necessity with the four modality markers across the varieties belonging to the outer and inner circle, first in the spoken dialogues and then in written press texts. This approach is useful in order to uncover the mechanisms underlying the asymmetries in their overall frequency of occurrence in the data (see Chapter 5), which might be related to different functional preferences. Statistical testing (Chi-square) will be employed wherever necessary to increase the explanatory power of the major contrasts in the data. Every modal instance occurring with each of the four markers was coded according to the established classification of either root or epistemic meaning (see section 4.6.1). To ensure functional equivalence between the various uses of these markers, semantic contrast was restricted to present tense, affirmative and declarative utterances. This is in keeping with the methodological steps set forth in section 4.3.3.

Figure 6.1, 6.2, 6.3 and 6.4 illustrate the root vs. epistemic distribution for *must, have to, (have) got to* and *need to* in the spoken dialogues under study (for a detailed overview of raw and normalized frequency counts, see Appendix 2.a). As expected, distribution of root modality in the present dataset exhibits an uneven pattern in ICE-JA and ICE-India, as compared to ICE-Ireland, ICE-GB and SBC, whereas epistemic necessity is less represented, especially in the former two varieties. Moreover, the figures show that the variant forms *must* and *have to* account for most alternation within root modality particularly in ICE-India and ICE-JA. In addition, ICE-JA exhibits functional layering with *need to*, whereas full layering of semantic contrast with all the four markers of strong obligation/necessity occurs only in ICE-Ireland. As already shown in the previous chapter, *(have) got to* has a marginal status in expressing semantic reading in ICE-JA and ICE-India (see also Figure 6.3), which contrasts sharply with the other three native varieties. It is ICE-Ireland which takes a middle position in this configuration.

However, the most striking contrasts are found in ICE-India, as the incidence of both root and epistemic modality is here the highest across all the corpora amounting to 53 and 37.2 instances per 100,000 words, respectively. Thus, the overall ordering of frequency counts is rather unevenly distributed (see also section 5.2) with Chi-square test (Yates' correction) yielding very high significance at $p<.0001$ and $p<.02$, respectively.[87]

[87] ICE-JA: ICE-GB Yates $\chi^2 = 51$, d.f. 1, $p<.0001$; ICE-IND: ICE-GB Yates $\chi^2 = 20.09$, d.f. 1, $p<.0001$; ICE-Ireland: ICE-GB Yates $\chi^2 = 0.01$, d.f. $p = .9203$(not significant); ICE-JA: SBC Yates $\chi^2 = 57.91$, d.f. 1, $p<0001$; ICE-IND: SBC Yates $\chi^2 = 29.16$, d.f. 1, $p<.0001$; ICE-Ireland: SBC Yates $\chi^2 = 5.07$, d.f. 1, $p<.0243$.

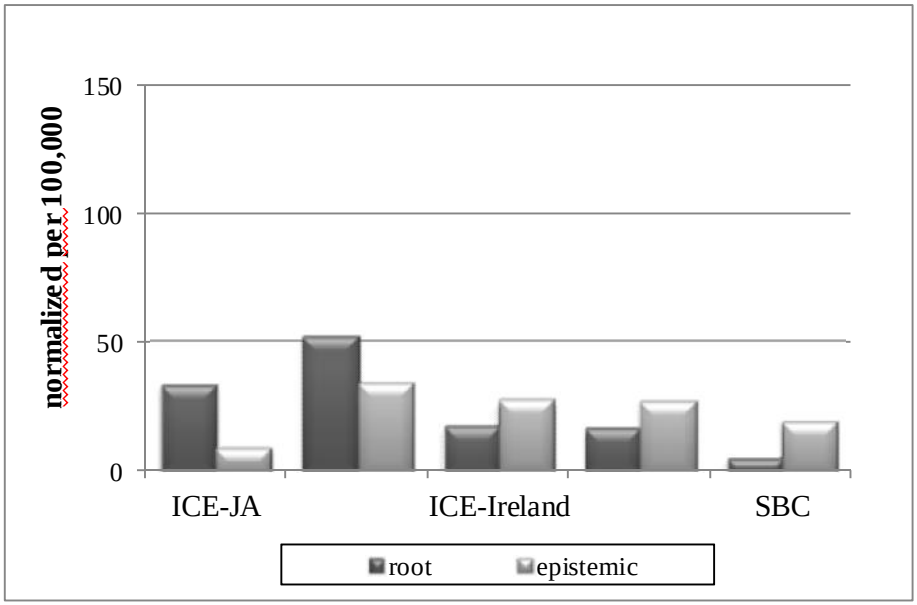

Figure 6.1. Distribution of root and epistemic meanings with *must* in per 100,000 words.

Furthermore, what the numbers in Figure 6.1 reveal, is that the root vs. epistemic distinction with modal *must* is also regionally displayed across the inner and outer circle. If evidence points towards a decrease of root and even epistemic necessity in BrE and AmE (Leech et al. 2009: 88-89), its preference can be confirmed in JamE and IndE (see Figure 6.1). By contrast, the data in ICE-Ireland, ICE-GB and SBC are more advanced in the use of epistemic *must*. As expected, epistemic uses outnumber the root *must* particularly in AmE. Incidentally, predominance of epistemic *must* appears to be a specific feature of native Englishes, which is also corroborated in previous corpus findings from AusE and NZE (Collins 2005). Thus, the initial hypothesis (see section 5.2.1) that the high frequency rate of this modal in the IndE dialogues correlates with high formality level and even with an authoritative style as the marker of a hierarchic society (Myhill 1995) might be taken as a relevant factor in explaining such differences in the data.

Preference of root obligation with *must* is not isolated to ICE-JA and ICE-India, but appears to also be common, for example, in SingE (e.g. Bao 2010) which could be explained by the specific contact ecology between Chinese or Malay – the two most often used native languages – and English. On the other hand, even though root uses prevail in the present outer circle varieties, the high incidence of epistemic *must* in ICE-

India is muchis more remarkable. Considering that epistemic modality is a later development in the modality system, it would be tempting to label IndE as more advanced than the other varieties. However, this is not supported by the high incidence of its root readings. Predictably, a higher level of formality is indicative of the fluctuations in the frequency of this modal in the IndE and JamE data.

By contrast, the low rates of root readings with this modal in the Irish data corroborates the evidence from non-standard Northern Irish English dialects (see Corrigan 2000; Tagliamonte & Smith 2006). Moreover, the IrE dialogues show many similarities with the figures for BrE both root and epistemic necessity.

In view of the findings from the previous chapter, it is less surprising that root *have to* is most common across all the corpora under study. Again, most instances of the semi-modal with this type of reading are overwhelmingly clustered together in the JamE and IndE conversations (see Figure 6.2) which are accompanied by high preference for root *must*. Only in the IrE dialogues does the normalized frequency of *have to* correspond more closely to the British data. The fact that the British dialogues show the lowest incidence of root readings with 46.9 instances per 100,000 words is rather surprising, but fits with the general trend emphasised in written data from BrE and AmE (Smith 2003: 257). As for the semantic contrast specifically in ICE-GB, the present evaluation on *must* and *have to* confirms to some extent the degree of variation found in Depraetere & Verhulst (2008: 17). Recall, however, that their study takes the whole spoken and written component of ICE-GB into account whereas the present study covers the use of these markers only in private and public dialogues. Furthermore, although the present author has followed the criteria for data selection closely, the analysis of *must* and *have to* in ICE-GB yields a higher number of epistemic readings as compared to their study.[88] By contrast, data from SBC (82.3 instances per 100,000 words) show almost full equivalence with ICE-Ireland in the distribution of root *have to* (80.2 instances per 100,000 words).

[88] A possible explanation for such asymmetry might be the fact that in some cases where the utterance did not provide sufficient context for interpretation the meaning was sought in the extended context of interaction. Likewise, the fact that the present analysis includes epistemic *must* as reporting anteriority (e.g. *You **must** have been a very fast driver* (ICE-GB.S1A-028.txt)) might also skew the statistics. This, however, is less often the case with *have to* (Collins 2009a: 64), and therefore the figures above should be interpreted with caution. For more details on epistemic necessity, the reader is referred to section 6.4.

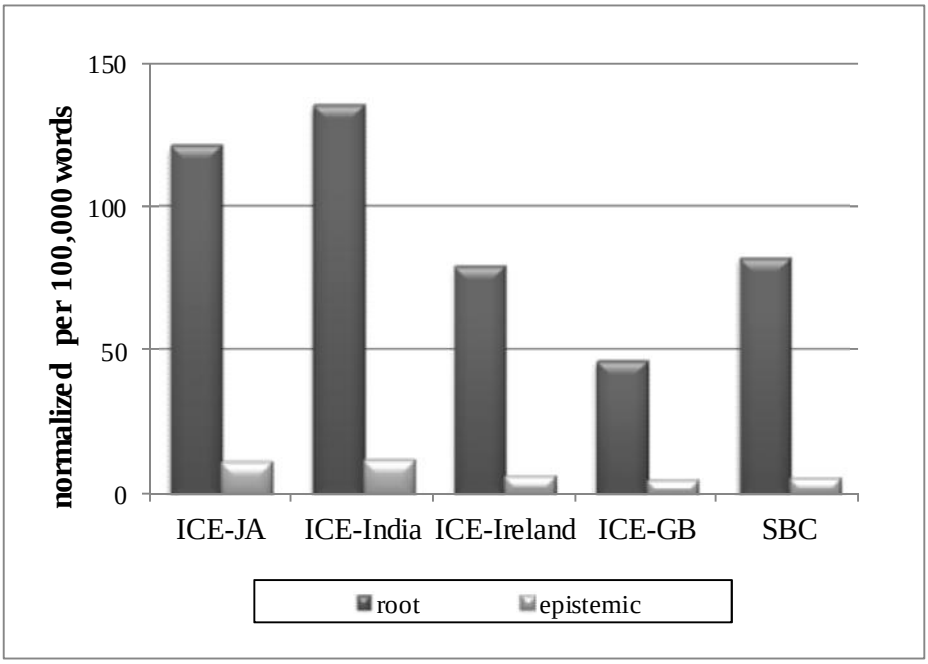

Figure 6.2. Distribution of root and epistemic meanings with *have to* per 100,000 words

The evaluation of epistemic necessity with *have to* shows a different ranking in the two outer circle varieties. As a recent development in present-day English, currently epistemic necessity with *have to* is considered to be still infrequent (Krug 2000: 90). Overall though the normalized frequency counts in Figure 6.2 are lowest for epistemic necessity, reinforcing the status of *have to* as a marker of predominantly root necessity, the figures for ICE-JA and ICE-India indicate a slightly higher rate of the former type of reading, which contrasts with the incidence found in the other three native varieties. Along these lines, in the previous chapter the high frequency rate of this semi-modal in ICE-JA and ICE-India was related to an earlier stage in development of obligation and necessity. One can only speculate that the same might be true with regard to the high incidence of root meanings. However, such an assumption would be counterintuitive in the case of epistemic readings, and, thus, two opposite tendencies in the expression of necessity in these two varieties can be seen which challenge the assumed colonial lag of the outer circle.[89] It is perhaps noteworthy to mention that such a pattern would correspond to the one in TorE found in Tagliamonte & D'Arcy (2007: 67-70). What is more relevant in this context is that similar to TorE, there seems to be a tendency of

[89] Cf. Corrigan (2000) on the development of epistemic *'be to'* in South Armagh English, which she accounts for as a colonial innovation.

reducing variability within root necessity in the outer circle varieties to favour *have to* particularly (Tagliamonte & D'Arcy 2007: 72).

The fact that epistemic *have to* is found at a lower rate in ICE-Ireland (with 6.9 instances per 100,000 words) compared to ICE-JA and ICE-India, but slightly higher than in ICE-GB (with 5.5 instances per 100,000 words) situates IrE closest to the American data from SBC (with 6.4 instances per 100,000 words). This is more remarkable as epistemic *have to* has been often considered a typical feature for AmE (Coates 1983: 57). As with modal *must*, from a cross-variety perspective the findings from ICE-Ireland for this semi-modal converge with those from non-standard data (see Tagliamonte & Smith 2006).

To sum up, the immediate impression when looking at the distribution of semantic contrast with *have to* is one of variation within and between the five varieties. However, statistical testing does not support such intuition as the values hardly proved significant ($p>.05$),[90] and therefore it could be explained by either random variation or further fine-grained distinctions which the frequency analysis of such a binary division cannot capture at this level.

As for *(have) got to* (see Figure 6.3), the bias in distribution of root and epistemic meanings is most spectacular in ICE-GB, SBC and ICE-Ireland. Due to the very low counts in the two outer circle varieties, a true comparison between the outer and inner circle seems bound to fail from the very beginning. Even so, the incidence of root necessity in ICE-India appears to be slightly higher than in ICE-JA.

[90] ICE-JA: ICE-GB Yates $\chi^2 = 0.35$, d.f. 1, $p=.5541$ (not significant); ICE-IND: ICE-GB Yates $\chi^2 = 0.68$, d.f. 1, $p=.4096$ (not significant); ICE-Ireland: ICE-GB Yates $\chi^2 = 0.68$, d.f. $p=.4096$ (not significant); ICE-JA: SBC Yates $\chi^2 = 0.28$, d.f. 1, $p=.5967$ (not significant); ICE-IND: SBC Yates $\chi^2 = 0.1$, d.f. 1, $p=.7518$ (not significant); ICE-Ireland: SBC Yates $\chi^2 = 0.03$, d.f. 1, $p=.8625$ (not significant).

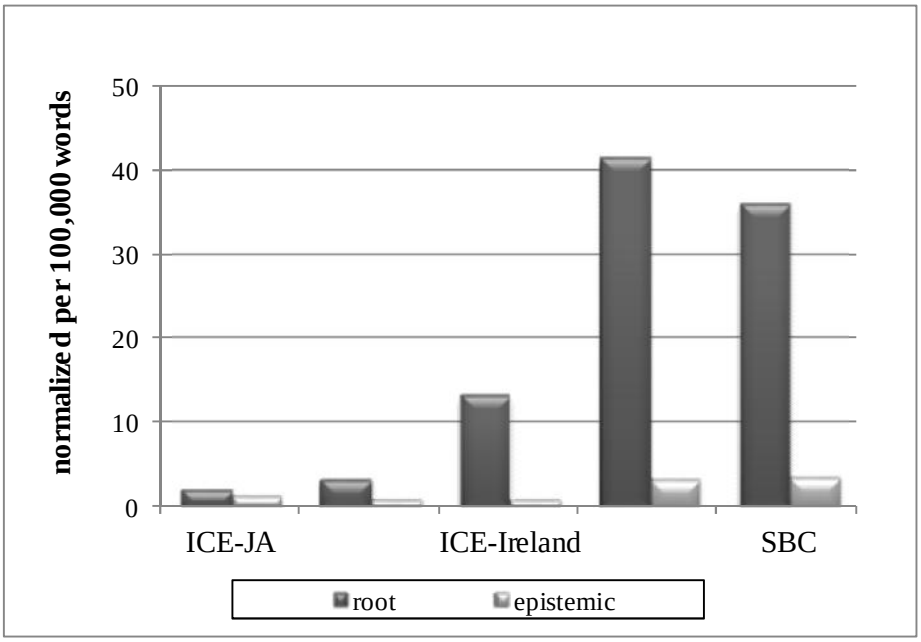

Figure 6.3. Distribution of root and epistemic meanings with *(have) got to* per 100,000 words

The stark contrast in the data is also statistically confirmed as it accounts for high significance at $p<.0016$ for the correlation between ICE-JA and ICE-GB, as well as at $p<.01$ between ICE-JA and SBC.[91] It is ICE-Ireland which shows consistent distribution of root *(have) got to* even though less than in ICE-GB and SBC. Such a functional distribution possibly mirrors the particular linguistic status of IrE as a younger native variety or merely the fact that, as in the non-standard IrE dialects (see Corrigan 2000), the use of this semi-modal is not that common.

[91] ICE-JA: ICE-GB Yates $\chi^2 = 13.2$, d.f. 1, $p<.0016$; ICE-IND: ICE-GB Yates $\chi^2 = 1.42$, d.f. 1, $p=.2334$ (not significant); ICE-Ireland: ICE-GB Yates $\chi^2 = 0$, d.f. $p=.1$ (not significant); ICE-JA: SBC Yates $\chi^2 = 6.58$, d.f. 1, $p<.0103$; ICE-IND: SBC Yates $\chi^2 = 0.69$, d.f. 1, $p=.4062$ (not significant); ICE-Ireland: SBC Yates $\chi^2 = 0.14$, d.f. 1, $p=.7083$ (not significant).

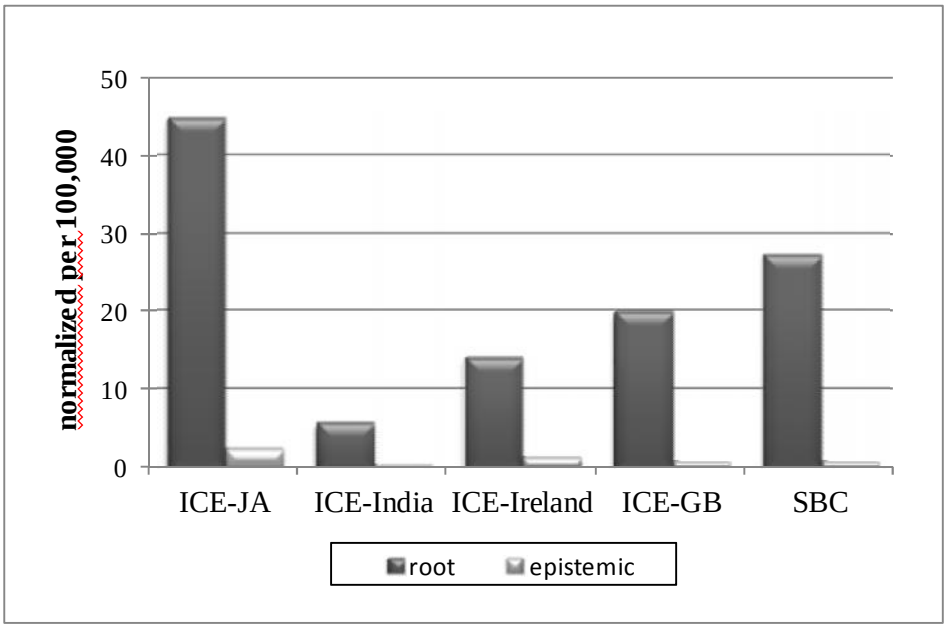

Figure 6.4. Distribution of root and epistemic meanings with *need to* per 100,000 words

Finally, semantic contrast with *need to* as the expression of unmarked necessity shows preference for root readings. Most occurrences are found in ICE-JA (46.3 instances per 100,000 words) followed by SBC (27.3 instances per 100,000 words) and ICE-GB (19.4 instances per 100,000 words). What is relevant for the discussion is that root *need to* in ICE-JA is more frequent than *must* (see Figure 6.1) entailing that Jamaican speakers have more options to express root necessity than the Indian ones. Although the corpus data displays a wide discrepancy in the distribution of semantic contrast with *need to*, statistical significance could not be proven in this case either.[92] Tentatively, it might be assumed that the high incidence of this recently developed semi-modal in ICE-JA appears to be more advanced also due to the time lag in compiling the Jamaican spoken component of the corpus (Mair 2009b: 19).

Regardless of the insufficient support to account for specific British or American patterns in the distribution of semantic contrast in the New Englishes, it will have become apparent that the data suggests various degrees of alternation between the root meanings of these markers. Before suggesting that the distribution in the spoken

[92] ICE-JA: ICE-GB Yates χ^2 = 001, d.f. 1, p=.9203 (not significant); ICE-IND: ICE-GB Yates χ^2 = 0.09, d.f. 1, p=.7642; ICE-Ireland: ICE-GB Yates χ^2 = 0.52, d.f. p=.4708 (not significant); ICE-JA: SBC Yates χ^2 = 0, d.f. 1, p=.1 (not significant); ICE-IND: SBC Yates χ^2 = 0.06, d.f. 1, p=.8065 (not significant); ICE-Ireland: SBC Yates χ^2 = 0.44, d.f. 1, p=.5071 (not significant).

dialogues might hinge on a combination between pragmatic and semantic constraints (see section 6.3), however, a comparison with the written press texts seems useful.

6.2.2 Written press texts

As with the case of the overall frequency patterns in the previous chapter which revealed a different ranking of items, semantic contrast in written press texts is less pronounced than in the spoken dialogues. Due to the partially inexistent or small number of epistemic readings, the present section will focus mainly on root necessity as summarized in Figure 6.5. However, for the detailed list of all the raw figures of the root vs. epistemic contrast, the reader is referred to consult Appendix 2.b.

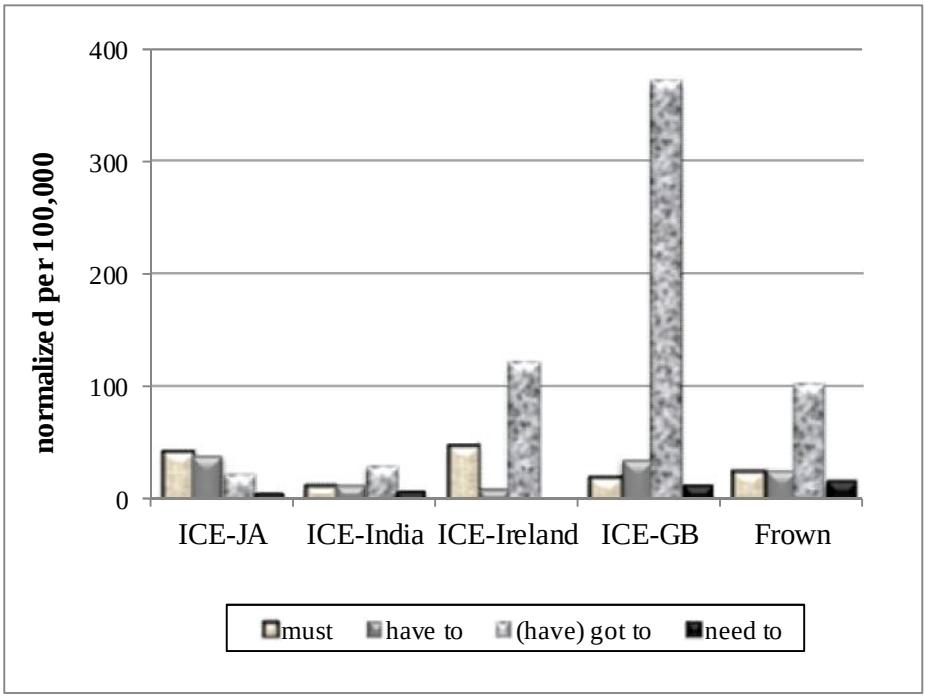

Figure 6.5. Distribution of root necessity in written press texts per 100,000 words

Suffice to say, epistemic readings occur sporadically with the modal *must*, whereas such readings with *have to* are confined to 1 occurrence in ICE-GB and 4 occurrences in SBC, respectively. Notably, epistemic readings with *(have) got to* are found in all the five datasets. Epistemic *need to* was not found in the data.

Despite the low incidence of root necessity, some interesting features are observable which contrast with the findings from the spoken dialogues. To a certain extent the distribution of root meanings replicates the patterns shown in Figure 5.4. Thus, the figures suggest that root *must* is most frequent in ICE-Ireland (with 47.5 occurrences per 100,000 words) and ICE-JA (with 42.5 occurrences per 100,000 words). It is ICE-India which reveals the lowest incidence of root meanings, whereas ICE-GB and SBC display a fairly balanced distribution. An evident aspect in Figure 6.5 which is worth mentioning is that root *have to* occurs rather marginally and seems to compete with modal *must* at least in ICE-JA and Frown whereas in ICE-GB the semi-modal outnumbers the modal. The most remarkable finding in the current context is that root readings of *(have) got to* outnumber all the other elements in the datasets representing native varieties, even in ICE-India whereas in ICE-JA this is a rather marginal element. Such a pattern differs significantly from the one in the spoken dialogues where semi-modal *have to* turned out to be the most often used root marker and *(have) got to* as the most infrequent one. Furthermore, it is worth mentioning that *need to* is missing in ICE-Ireland. While such differences would invalidate the relevancy of the previous findings, recall that the functional analysis of obligation/necessity in the written press texts considered only present tense, affirmative and declarative utterances, as this is also the context which most adequately accounts for alternation between variant forms. When compared with the patterns from Figure 5.4, the functional distribution tentatively suggests that the press texts in the outer circle varieties exhibit a more formal profile and thus, they follow more 'educated' conventions of written English. This is, however, contrasted by the evidence on epistemic *must* and *(have) got to* alternating in the press texts in all the corpora under study, though, admittedly at a very low level of discourse frequency.

To conclude, the press texts in the other three varieties confirm the previous claims about the stylistic shifts in written language towards more colloquial and informal features, and which suggest that journalistic prose is among the most progressive genres at least in BrE and AmE. Since both the press texts under analysis and the occurrences are quite small in number, further research on larger datasets will be necessary to uncover these tendencies.

6.3 Root modality: description of sources in the New Englishes

Even though the examination of distribution patterns in the spoken dialogues from section 6.2 did not yield any conclusive answer for a clear British or American orientation in usage, the present data suggests vigorous alternation at least between root *must* and *have to* in ICE-JA, ICE-India and ICE-Ireland. To these two, the semi-modal *need to* should be added as competing with the other two items in ICE-JA and *(have) got to* in ICE-Ireland. As observed in other native varieties of English, e.g. in CanE (Tagliamonte & D'Arcy 2007: 63), such alternations might involve a reorganization of root necessity in these varieties, or individual items might be constrained to contexts which convey functional extension or specialization. Therefore, a more fine-grained qualitative perspective is necessary to disentangle the various relations in the use of these items.

Although previous descriptions on present-day English have admitted that both *have to* (Collins 2009a: 61; Depraetere & Verhulst 2008; Westney 1995: 118) and *need to* (Nokkonen 2006; Smith 2003; Taeymans 2004) potentially involve subjectivity such distinctions have been hardly examined in data from New Englishes. Therefore, to assess the potential overlap in the root readings of these markers the ensuing examination focuses on the subjective vs. objective dichotomy in data from ICE-JA, ICE-India and ICE-Ireland (see classification in section 4.6.3). Furthermore, while the assessment of subjectivity in the use of root necessity hinges on pragmatic inferences, register and level of formality will be considered as factors constraining the source of necessity (Depraetere & Verhulst 2008: 24). All in all, the present analysis takes into account the variable context which accounts for the choice of one meaning over the other.

Increased subjectivity is a common tendency which has contributed to the subjectification of modal markers in present-day English. The conventionalization of uses through subjectification is accepted for *must* and has already been accounted to some extent for the extension of obligation *have to* to epistemic necessity (Krug 2000: 61; Traugott 1988: 411). Thus, the aim is to assess whether such contrasts can function as diagnosis for ongoing semantic change *within* root necessity, which would implicitly

account for a rearrangement in the domain of strong obligation/necessity in standard JamE, IndE and IrE.

In view of the diachronic development of modals as root/deontic – to – epistemic meanings, central to the analysis is a narrower path of semantic change, which is schematized in Figure 6.6:

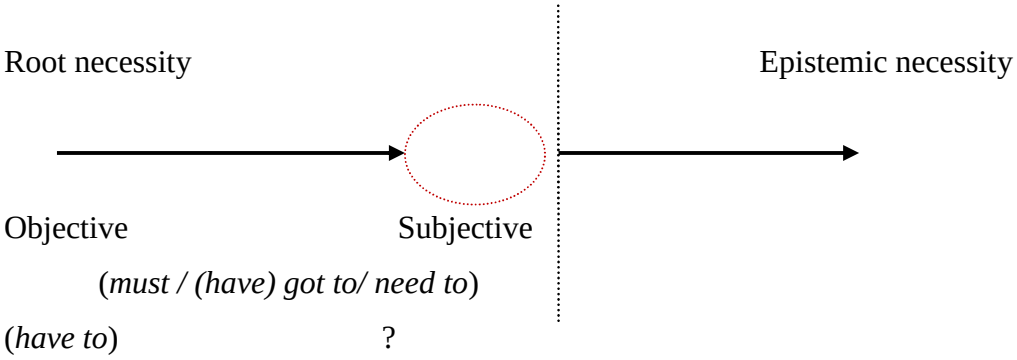

Figure. 6.6. Semantic shift from objective to subjective modality in New Englishes

The figure above illustrates the specific shift from objective to more subjectified meanings within root modality that accompanies the structural changes in the grammaticalization of these items (see Goossens 2000; Krug 2000; Nordlinger & Traugott 1997; Traugott & Dasher 2002), and which is captured schematically with the circled dotted arrow. Thus, from the perspective of previous accounts on this grammatical area, the modal *must* and the semi-modals *(have) got to* and, more recently, *need to* occupy an intermediate position on this cline of development, which point to their ability to be used both objectively and subjectively. It is the semi-modal *have to* which arguably expresses only objective necessity.

The discussion focuses both on the most typical and infrequent root realizations with *must*, *have to*, *(have) got to* and *need* with relevant examples from the data. However, increased attention is given to the alternation between the first two items which will help draw the modal profile of strong obligation and necessity for each variety. The section ends with a discussion of the main patterns of use identified in each variety. Although such an analytic approach is essentially synchronic, a central issue is how to accommodate the theoretical assumptions about semantic change as progression from objective to subjective necessity of root readings in New Englishes. The analysis shows that variation among the four modal markers in spoken interaction from ex-

colonial Englishes, is displayed not only quantitatively, but reflects subtle differences in meaning conveyed by the root vs. epistemic duality, which contrasts with BrE and AmE.

6.3.1. Discourse-external sources: rules and regulations

6.3.1.1 Must

Considering the findings so far, what is then the motivation that primarily underlies the alternation between *must* and *have to* in the database, and why are root readings with the modal so wide spread in ICE-JA and ICE-India considering the general decreasing trend in BrE and AmE? And, is the widespread use of root *must* in ICE-JA and ICE-India a case of colonial lag, of "recessive innovation" (Hundt 2009) or "convergence-to-subtratum" (Bao 2010) or just variation? The ensuing analysis reveals that modal *must* in spoken data from JamE, IndE and IrE communicates a wider range of sources of root readings which go beyond the speaker's imposition of obligations on the addressee to complete an action in that it occurs also non-performatively. This wider range of sources builds on the possibility stated in Depraetere & Verhulst (2008: 3) that the speaker might be only reporting the necessity, and thus is the 'channel' or that, as Myhill (1995: 169) notes, *must* involves an emotional reaction to obligation with the event usually being motivated by abstract social norms.[93]

Arguably, the obligation is particularly strong[94] if the subject preceding the modal is second person *you*. There seems to be a link between source of necessity and the subject of the utterance whether it is in the first, second or third person as argued in Coates (1983: 34-35), which will be explored in the present database. From this perspective, despite the strong speaker-oriented obligation in utterance (67) extracted from a conversation between three men aged 26-33, the source for the motivation of an Indian to get married is dictated by a social rule such as reaching a sufficiently mature age – which, admittedly, may equally function as a condition:

[93] However, for Myhill (1995: 179) it is in fact *got to* which genuinely involves emotional motivation of root modality, and which is different from "principled" *must*.

[94] In this context 'strong obligation' refers to pragmatic strength, i.e. pragmatic strengthening or weakening of the modal meaning in an utterance (cf. Depraetere & Verhulst 2008: 15, footnote 23). See also Huddleston & Pullum (2002: 176).

(67) Uh no you **must** get married because now you are twenty-six or twenty-seven <w> it's </w> time to marry (ICE-India.S1A-024.txt)

A more careful reading of this example shows the difficulty in gauging the objective nature of the source, as the necessity of getting married might be attributed to the speaker's personal wish as the authority invoking social norms (cf. Depraetere & Verhulst 2008: 6). Although such an interpretation emphasises the complexity of root modality, the reported obligation in the instance above is more a comment on the addressee's statement of not wanting to get married. By contrast, the second pronoun *you* preceding modal *must* in example (68) from ICE-Ireland expresses an external necessity implicitly related to the institutional rules in the case of a grant application:

(68) And you **must** send the form that they send off with the letter saying that they 're giving you the grant <#> You **'ve got to** send that back within a week (ICE-Ireland.S1A-053.txt)

More remarkable is the use of contracted *have got to* (e.g. '*ve got to*) in the following conversational turn which contrasts with *must* and seems to suggest an obligation linked to an "urgent emotional reaction" (Myhill 1995: 167). While the choice of modal *must* would be perhaps more appropriate, it is the urgency of the external necessity in this case, i.e. to send the application within the requested deadline, which the speaker emphasises here. Conversely, semi-modal *have to* would be an equally good choice in this utterance given that it may report on specific regulations, but distinguishes itself from the other two markers through overt neutrality. Incidentally, such a choice appears to be supported by generic *you*. Within the IrE context there is another aspect which makes utterance (68) a noteworthy. Regardless of the fact that the high discourse frequency of *have to* outnumbers the other markers altogether, it is also the informal context of a face-to-face conversation which might be invoked in the choice for contracted variant '*ve got to*.

Similarly, generic *you* preceding *must* in excerpt (69) signals objectivity in that the source of necessity cannot be identified with the speaker – although it might be argued that the speaker imposes the necessity on them – but rather stems from the specific situation which refers to the correct method of breathing:

(69) <#>The breathing is a challenge for me and I remember while learning to swim in Florida State University they taught me to breathe through my nose

<{11><[11>and</[11> somebody told me that that's for competitive swimming <#>But now I'm learning that you **must** breathe <{12><[12>through your mouth</[12> (ICE-JA.S1A-020.txt)

In this sense, modal *must* does not impose an obligation, rather its use is based on a natural reaction of the body which converts itself into a necessity, and that in case of non-compliance under certain circumstances one might have to support the consequences (cf. Bybee et al. 1994: 186). In the present context, however, the circumstances denoting necessity are such that the use of semi-modal *have to* in the present tense would emphasise the existence of a general or repeated requirement, e.g. learning how to breathe, which is not linked to the speaker's imposition (Westney 1995: 112). Considering that root meanings with modal *must* are associated with some binding force (e.g. social norms, rules, traditions etc.), it might be perhaps the urgency of the situation that the speaker wishes to emphasise here.

External root meaning with modal *must* may co-occur with a first person subject pronoun as in (70) where the source of necessity is conveyed by the specific situation of someone not understanding English and instead using Patois:

(70) And here we are now saying they've come to you as a government officer and they don't understand English so we **must** speak to them in <{7><[7>Patois <#> (ICE-JA.S1A-040.txt)

Although the necessity here appears to be assigned by the first person plural subject pronoun *we* –which could be interpreted as indicating an internally motivated necessity – the source of necessity emerges from the specific circumstance when a Jamaican person interacts in Patois with a civil servant because of his/her poor command of English. Replacement through *have to* in this case would in fact emphasise the objective origin of necessity so that encoding of root meaning is least affected by indeterminacy. On the other hand, the modal appears as a more appropriate choice in this context if the seriousness of the situation presented in this conversational turn is taken into account, and which compels the persons involved in it to act contrary to their customs.

Inescapable necessity motivated by law using modal *must* as in (71) suggests that non-compliance might have severe legal consequences on the individual (cf. Depraetere & Verhulst 2008: 15 f):

(71) <#>Everybody **must** comply because you once it's in the Constitution and I don't speak to you in Patois you know <{10><[10>you can take me to court</[10><,> (ICE-JA.S1A-038.txt)

At first glance, the indefinite pronoun *everybody* appears to weaken the effects of non-compliance to speak in Patois, the subordinate in fact strengthens it by alternating pronoun subject *I* and *you*, which ultimately increases the pragmatic force (i.e. legal consequences) of the utterance. In contrast to institutionalized rules and regulations, the use of *must* to express necessity as linked to universally accepted social norms is less compelling. Thus, in (72) the necessity originates in the principles of proper behaviour in life:

(72) It teaches the moral that one **must** work hard and save for the future <,,> (ICE-India.S1B-012.txt)

Again, the subject, here impersonal *one*, is indicative of the weak impact of the necessity when the speaker's choice falls on the modal. Here the speaker, also the channel of the reported necessity, explores the principles of conventionalized behaviour in life on the basis of the famous fable of the ant and the grasshopper. As this example is part of the text category 'class lessons' the use of *must* in this context is exemplary for an intentional educational/moralizing purpose of the speaker. In this context, the modal expresses an "irresistible force" (Sweetser 1990: 54) which resembles an inescapable situation to which everyone is bound to comply. A similar usage is found in instance (73) which is extracted from the same text category of 'class lessons' from ICE-India:

(73) To see the universal and all pervading spirit of truth face to face one **must** be able to love the <,> meanest of creation as <,> oneself <,,>(ICE-India.S1B-011.txt)

In this example, it is not very clear where the source of necessity derives from, as it might as well be the speaker urging the addressee to show affection, which in this context appears to be a rather odd formulation. However, it becomes immediately apparent that the specific circumstance in this case is a condition, which lies at the heart of seeing the spirit of truth, and that there is no human control as a higher force which is

involved.[95] It can be also interpreted as a necessity from the addressee's perspective – although rendered here with generic subject *one* – with the modal increasing the moral commitment of the agent, and if it were replaced with semi-modal *have to* the effect of necessity would be weakened (cf. Westney 1995: 117). While in present-day English such uses would be interpreted as "unduly insistent" (Smith 2003: 259), they appear to be a common choice in the spoken dialogues from ICE-India.

Undoubtedly, any of the other four markers would appear to be an appropriate choice in this utterance, barely affecting the realization of the root meaning.[96] Thus, *have got to* conveys urgency to the utterance, whereas with *need to* the utterance acquires a different motivation as internal to the speaker and at the same time minimizes the effect of the necessity. All these examples are very close to the original root meaning with *must* as associated to an external circumstance, e.g. a condition or some higher force and which has developed to express speaker imposition (Traugott & Dasher 2002: 135). Moreover, in both (72) and (73) the external root meaning of *must* makes an assertion about the whole proposition emphasising the actualisation of the event, and thus has wide scope (Nordlinger & Traugott 1997: 302; Traugott & Dasher 2002: 130). Such uses demonstrate that newer and older functional layers of the strong obligation/necessity have persisted along with the formal distinctions found in New Englishes. In regards to these developments but bearing in mind the prior frequency analyses (see Chapter 5 and section 6.2), it is perhaps not surprising that such uses occur particularly in the category of IndE class lessons.

Specific circumstances are common discourse-external sources for a necessity to be fulfilled. In example (74) it is the agreement on rules which is the condition underlying the initiation of a discussion on the problems in Northern Ireland:

(74) We wish to embark upon a process of discussion among constitutional parties to resolve the difficulty in Northern Ireland <#> But before you embark upon a consultative process and discussion we **must** agree the rules of discussion <#> (ICE-Ireland.S1B-022.txt)

The specific condition in this example is worth mentioning as it is introduced by first person subject pronoun *we* and resembles more of an 'inner urge' (Depraetere &

[95] Cf. also Westney's (1995: 112) interpretation of *must* as displaying an "unqualified requirement".
[96] On the changes in the effect of modal meaning between variant forms of necessity, see also Palmer (1990 [1979]: 115).

Verhulst 2008: 8) in order to avoid unpleasant consequences or sanctions. Despite the alternation between generic second person *you* and definite first person *we* as the agent of the necessity to comply with rules, speaker involvement is minimal. It is the implicit condition which refers to the actualization of the situation of starting a consultative process and discussion, which, in addition, increases the force of the modalized utterance. Clearly, *must* best suits contexts linked to institutionalized obligations and necessities, and which correspond with the diachronic development of root meanings. In addition, it is also the formal context of the text category 'broadcast discussions' which favours the modal, as such uses are also more typical for the written style. Consider for comparison example (75) selected from the written press texts in ICE-Ireland, where the necessity lacks speaker involvement and originates in the conditional clause (i.e. *if he is to meet a deadline…):*

(75) The Irish diplomat **must** win new concessions from the hardline French government if he is to meet a deadline for a deal in December. (ICE-Ireland.W2C-008.txt)

While discourse-external root meanings with modal *must* are considered to be more common in written language, the data shows that such uses are often employed in formal contexts of spontaneous spoken interaction, as in (77), (78), (79) and (80):

(77) Uh which means that <,> the micro-organisms by themselves <,> cannot fix nitrogen but they **must** <,> uh enter into a symbiotic arrangement <,> with a particular plant <,> (ICE-India.S1B-046.txt)

(78) They the government attempted to justify their surrender to terrorism <.> b </.> <,> terrorism by telling us that the <.> par </.> paramilitary parties would only gain entry on the foot of an absolutely permanent ceasefire <#> They **must** hand over their illegal weapons <,> we were told <#> But as the months and years pass by this Government have watered these conditions down until they now will welcome the IRA to talks complete with their deadly arsenal <#> (ICE-Ireland.S1B-053.txt)

(79) Thirdly <,> he says a declaration by the two governments that Sinn Fe/in 's entry into the talks at any stage **must** be preceded by the declaration of a complete and permanent ceasefire <#> And indeed <,> insofar as a government 's declaration is worth anything that (ICE-Ireland.S1B-053.txt)

(80) Developing countries **must** consent to pursue reforms <#> (ICE-JA.S1B-018.txt)

In (77) the necessity expressed with *must* refers to the particular circumstance of entering into a symbiotic arrangement which cause a reaction within micro-organisms; in (78) and (79) it is the condition of absolute ceasefire which underlies the necessity of handing over the illegal weapons or of a declaration of the two governments, respectively; finally, in (80) the necessity is dictated by the extended context of the utterance, which is the consent to initiate reforms in Jamaica in order to eradicate corruption with the ex-colonial constabularies in Jamaica. Notably, the most neutral meaning in the sense of non-relatedness with the speaker is found in example (77) which is part of a broadcast discussion on the importance of research in biology and genetic engineering in India. The modal here is devoid of any emotional or urgent obligation as it is related to explaining a specific biochemical process (i.e. fixing nitrogen) with the outcome of necessity as bound to the laws of nature. Furthermore, it is the formal context of the text category which might be invoked in the choice of *must*. By contrast, the necessity expressed in (78) and (79) is directly linked to urgency of an inescapable situation based on strong, in this case, moral and political principles typically employed in 'parliamentary debates', and which therefore require modal *must*. Example (80) is similar in this sense as the consent to pursue reforms is considered an indisputable condition to the progress of developing countries. Furthermore, the fact that this utterance is part of the text category 'class lessons' increases the motivation for choosing modal *must*.

In addition, there are specific circumstances or conditions which reduce speaker involvement. Unless the speaker is the channel reporting the necessity, such necessities do not typically co-occur with an external source. Excerpt (81) refers therefore to a specific condition which is necessary for the speaker to perform a joke:

(81) <S1A-081$A><#> That 's too low <#> Right what 's the joke
<S1A-081$E><#> I 'll tell ye later on <&> clears throat </&>
<S1A-081$A> <#> Uh
<S1A-081$B> <#> The audience **must** be right (ICE-Ireland.S1A-081.txt)

It might be argued that the specific condition here derives from the speaker's internal disposition interpreted as a face-saving strategy for successful communication. However, since in general one cannot influence the reactions of an audience, it's unclear whether there is an imposed obligation. Therefore, the necessary condition in this

example has an external source which does not identify with the speaker. Similar to (73), the source of necessity in (82) could be interpreted as internally imposed on the subject as the agent of necessity, which is here referred to as *the modern man*:

(82) The modern man nowadays must have <,> **must** be jangling a bunch of keys <#> Keys of what <#> The keys <,> of a motor car <#> A BMW (ICE-Ireland.S1A-061.txt)

Following Depraetere & Verhulst (cf. 2008: 8), even though the necessity of *jangling a bunch of keys*, appears to be the object of the internal disposition of the subject, it is the particular context which justifies a discourse-external interpretation. In this sense, the subject, i.e. *the modern man*, is the agent which complies with his own disposition as identified in the extended context of the utterance. Interestingly, such realizations of root necessity are more common with *need to* (see section 6.3.1.4). However if the modal were replaced, the meaning would lose urgency, as the necessity in this sense appears to be an uncontrollable urge for the subject to react in a specific way. In fact, a more appropriate choice would be *(have) got to* which enhances the effect of urgency. Such uses where the subject-internal necessity is accompanied with external factors might be seen as a transition towards more subjectified meanings of this modal.[97]

6.3.1.2 Have to

Unlike modal *must* which is most often associated with a subjective interpretation of root necessity, discourse-external sources are more common with semi-modal *have to*. Although the use of *have to* is associated with a general requirement or neutral necessity, the description on the basis of discourse-external sources reveals that, similar to *must*, root necessity in other native and non-native varieties of English covers a wide range of realizations. As with *must*, an additional factor involved in the modalization with *have to* refers to level of formality, i.e. formal vs. informal contexts of interaction.

It was argued earlier how modal *must* with a discourse-external source is often related to rules and regulations and the modality expressed does not involve human control. Semi-modal *have to* in example (83) representing a face-to-face interaction

[97] Exemplary in this sense is Goossens' (2000: 157-159) analysis on the partial chaining of OE *magan* even though the shift is reversed from participant-internal to participant-external necessity.

from ICE-JA is an example of such a source, which appears to be quite common in casual conversation:

(83) That just assumed that once you get on the air there and now in mass communication then you **have to** speak the language of <,> the official <}> <-> <.> la </.> </-> <=> language </=> </}> <{1> <[1> <,> </[1> which is English <{2> <[2> <,> </[2> <#> (ICE-JA.S1A-001.txt)

Here, generic subject *you* emphasises the degree of necessity as related to an external circumstance, which in this case refers to a rule on the use of the official language. Thus the source lies outside the imposition of the speaker. In this sense, semi-modal *have to* contrasts with modal *must* and appears to be a more appropriate choice to indicate the existence of some authority, but neutralizing the effect of the necessity. As with modal *must*, strong requirement can also be expressed with *have to* as in the following instances (84) to (88) which are extracted also from the private dialogues:

(84) You **have to** come in by a certain time <#>So you **have to** abide by certain <{2><[2>rules</[2> <$A> (ICE-JA.S1A-024.txt)

(85) Uhm no in this case <,> here as per the rules he **has to** <{><[> pay </[> fine (ICE-India.S1A-018.txt)

(86) Well of course we **have to** speak English I suppose eh cos your interest is English <$A><#> (ICE-JA.S1A-084.txt)

(87) He **has to** get the work done cos he will be removed (ICE-JA.S1A-008.txt)

(88) Everybody **have to** go to the chapel <#> (ICE-JA.S1A-056.txt)

While example (84), (85), (87) and (88) are relatively unproblematic in identifying a regulation as source of the necessity– i.e. regulations on the university campus in the first, the obligation of paying a fine in the second, or to fulfil the work load in the fourth, and an imposed rule in the last – the necessity in (86) is noteworthy. The objectivity of this utterance emerges from a situation which imposes a certain action, such as speaking in English because the focus is on English in the conversation. From a pragmatic point of view though, the semi-modal expresses agreement in this case. On the other hand, it might be argued that the semi-modal here is covering a wider scope of reading almost in the sense of a conclusion which is characteristic of logical necessity,

and which Coates (1983) has classified as *merger*. Conversely, the modal *must* in this context would increase the existence of a binding force (Westney 1995: 118).

Likewise, in (89) to (92) the necessity has an external motivation which is independent from the speaker, thus giving the utterance a generalized meaning as 'unmarked' (Leech et al. 2009: 110). Even though first person pronoun *I* appears to be a strong indication of a subjective source, a particular circumstance, i.e. such as cooking, is most likely the source of necessity in (89):

(89) Yeah and so I **have to** go home and cook anyways<#> (ICE-JA.S1A-075.txt)

(90) I **have to** go to Madurai <,> (ICE-India.S1A-024.txt)

(91) Well <,> I **have to** go home (ICE-Ireland.S1A-093.txt)

(92) **Have to** be here on Ash Wednesday to work (ICE-JA.S1B-043.txt)

Even though an objective root interpretation for example (90) and (91) is most suitable, the necessity is not overtly marked as external to the speaker, and seems to more closely resemble the semantics of *need to*. There might be a hidden speaker intention of an external necessity though, to go to Madras in (90) and to go home in (91).

A further particularity is subject ellipsis as found in (92) which is part of a 'broadcast interview' from the Jamaican component of public dialogues where this speaker complains about the working conditions of librarians. As it has been already suggested, *have to* usually refers to a general requirement. However, this example is noteworthy as the speaker in fact emphasises the actualisation of a specific circumstance. Both *must* and *(have) got to* would be possible in this utterance, with the latter item increasing the urgency of the external circumstance.

More typical in the case of *have to* are root readings related to habitual actions, in this sense the present database provides several examples, such as (93) and (94):

(93) Oh I **have to** go to work</quote> like Jamaicans like <quote>ah I *have to* go to work this morning</quote> and they don't want to come out of bed is like sleepy (ICE-JA-S1A.008.txt)

(94) And that's good <{1><[1>cos he **have to**</[1> do his work you know<{2><[2><,></[2> <#> (ICE-JA.S1A-008.txt)

Similar to modal *must*, necessity as related to the internal disposition of the speaker may also occur with the semi-modal *have to* as in (95) where the action of coming in is dictated by the internal needs of the agent, in this case definite pronoun *they*:

(95) And they **have to** come in and they'll probably watch T V until two o'clock I don't know once they go out they can't come back in unless again bathe or wash <?>their feet</?><O>laughter</O> (ICE-JA.S1A-011.txt)

As the third person subject pronoun also indicates, there is no authority in charge imposing the necessity. Interestingly, such use of *have to* corresponds to that of *must* as exemplified in (82) with the necessity motivated by an internal disposition of the agent which is at the same time linked to an external circumstance or condition. From this perspective the use of the semi-modal *have to* in such a context is quite odd. Nevertheless, this example also shows that other markers of necessity have started competing with *need to*.

Pragmatic weakening with *have to* in the sense of an existing necessity is evident in (96), where the source refers to a requirement which is not binding to the agent and thus may be considered as optional:

(96) So as a consultant <}><->you</-> <=>you</=></}> speak English uhm but you get
different flavours of understanding<O>laughter</O> right so<{3><[3><,></[3> you **have to** speed up slow down<{4><[4><,></[4> depending on how the people are in terms of how you deliver your message (ICE-JA.S1A-014.txt)

Likewise, objective or neutral necessity with this semi-modal often occurs in the public dialogues as illustrated below from (97) to (100):

(97) But in the case of these animals <,> they **have to** depend on the plants <,> for the amino acid source (ICE-India.S1B-001.txt)

(98) <ICE-IND:S1B-010#192:1:A> When they form ice it requires more places <,> and so therefore the rock <}> <-> is <,,> has to get </-> <+> gets </+> </}> expanded <,> <ICE-IND:S1B-010#193:1:A> The rock **has to** get expanded <,> okay (ICE-India.S1B-010.txt)

(99) So consequently you **have to** do institutions that are left in this business which is the Building Society the National Housing Trust and the Jamaican Mortgage Bank are not<,> and will not do financing for units that are sold for more than three to four million dollars (ICE-JA.S1B-029.txt)

(100) But you **have to** make a distinction between the management of the organisation and the execution of certain statutory functions which are governed by legislation (ICE-Ireland.S1B-068.txt)

Both (97) and (98) refer to an existing necessity which underlies the laws of nature which does not involve human control. The two examples which are part of the Indian text category of 'class lessons' are in clear contrast with (77) where the modal *must* was used. In that particular example I argued that the modal is used as a marker of neutral necessity which in fact is more typical with *have to*. From this perspective, the semi-modal is a more appropriate choice in such contexts.

The necessity in (99) and (100) is also objective as it appears to be used as a rhetoric device. Generic second person *you* in (99) is an indicator for such interpretation with the speaker merely emphasising a general requirement. Likewise, generic *you* in (100) is not related to speaker authority as it occurs almost in an idiomatic sense which enhances the interactive context of conversation.

6.3.1.3 (Have) got to

Due to the informal colloquial connotation attached to *(have) got to*, this construction has been often equalled in meaning with *have to* (e.g. Coates 1983; Palmer 1990 [1979]). The following examples, (101) to (107), relate the use of *(have) got to* with an external source of necessity, however, showing a difference in the strength of meaning than *have to*:

(101) She'**s** </w> **got to** work from <,,> nine <w> O'clock </w> till five (ICE-India.S1A-021.txt)

(102) ICE-IND:S1A-055#258:1:B> No you *have to* wait for your <,> marks sheets <,> whether you passed or fail <,> <$A> <ICE-IND:S1A-055#259:1:A> Oh you **have got to** wait for your marks first (ICE-India.S1A-055.txt)

(103) And once you do that you **got to** be the saviour of the institutions <w> You'**ve** </w> **got to** be <,> promoting causes <w> (ICE-India.S1B-050.txt)

(104) You'**ve** </w> **got** be doing things <,> which earlier you <w>didn't </w> (ICE-India.S1B-050.txt)

(105) And if the bus system is to be continued to be operated at that level<{4><[4><,></[4> then its economic costs **have got to** be met (ICE-JA.S1A-093.txt)

(106) There there 's something in Paul that feels that he **has got to** save the world and the world being this whole political process (ICE-Ireland.S1A-043.txt)

(107) > Yeah <,> so I get to keep the bar by myself <#> But the only thing is that uhm <,> sometimes it can get really crowded in there and you **'ve got to** clear the tables and serve at the bar and everything all at the same time like (ICE-Ireland.S1A-057.txt)

While (101) states a necessity related to a rule or a fixed working schedule, the speaker of this utterance appears also to add more emphasis in the sense of 'circumstances compel' (Palmer 1990 [1979]: 114). Intuitively, such meaning appears to be closer to what was earlier mentioned regarding modal *must* in (72) and (73). Similarly, despite the external sources in the following examples, the necessity here is more emphatic than in the cases with *must* and *have to*. In addition, the external factor is underlined by the third person pronoun *she.* Interestingly, such uses are more common with the semi-modal *have to,* whereas *must* would increase the compelling meaning. Somewhat different in meaning appears to be (102), where the speaker uses '*ve got to* as an acknowledgement or response to the previous speech turn. Although logical necessity might be invoked in such cases, the necessity in fact is externally motivated by a specific circumstance which compels the agent to wait for the marks. It follows from this that the semi-modal is imbued with an interactive meaning. Somewhat similar to the first example, (105) reports on a condition which underlies the necessity of an improved transportation system. While such a meaning is closer to a specific requirement, it is also compatible with the stylistic neutrality of *have to* as well as with a dynamic interpretation of modal necessity (Collins 2009a: 68, 70). Alternation with *have to* is also nicely captured in (102), where speaker B uses *have to* to convey objectivity to the particular condition whereas speaker A, as mentioned earlier, uses variant '*ve got to* to reinforce the urgency of such a situation. From these examples it becomes apparent that this semi-modal is slightly different than *have to*.

As in (82), the variant forms *got to* and '*ve got to* in (103) and (105) there is no human control in the necessity reported as the source here is external to the speaker, but which constraints the agent (i.e. second person pronoun *you* and third person *he*) to

embrace the situation and react accordingly. More remarkable is that the former utterance occurs in a 'broadcast discussion'. Therefore it can tentatively be asserted that *(have) got to* supplies the necessity with an interactive function. Thus, example (104) is akin to the previous one. In fact, it is the continuation of the previous speech sequence where the speaker asserts the urgency of the specific situation, i.e. of doing things which one did not do before, which resembles an insistent demand on the agent, here generic second person *you*. Perhaps even closer to the necessity expressed in (82) is (106), which is equally extracted from the private dialogues in the Irish sub-corpora. Again, the necessity emerges from the speaker that reports on the intrinsic disposition of the agent urging him (i.e. *Paul*) to save the world. In contrast with (82), the choice for *(have) got to* in such contexts is more appropriate as it underlines the urgency of the circumstance. Predictably, such uses are borderline cases of objective root necessity to more subjective meanings.

Finally, instance (107) relates the necessity to a general requirement (circumstance) as is more typical with *have to* and which is also reflected in the use of generic second person pronoun *you*. On the other hand, the urgency expressed with this semi-modal emerges from the specific situations expressed in the subordinate, i.e. *sometimes it can get really crowded in there*.

Obligation with *(have) got to* pointing towards external sources arguably may have the same strong effect as with modal *must*. As it is apparent with discourse-internal sources (see 6.3.2.3), such uses are quite common in public statements as found in (108):

(108) I must say to this House today that no-one who has committed a crime can <,> have an amnesty <#> They **have got to** be tried for their crime (ICE-Ireland.S1B-051.txt)

This instance occurs in a public context of interaction as found in 'parliamentary debates' from ICE-Ireland. Although this utterance is a strong directive, the objectivity emerges from the extended context which refers to the binding force of moral principles and legal laws, that any crime must be condemned. Thus, the speaker is only the channel which emphasises the inescapable obligation to which everyone is forced to abide (cf. the example with modal *must* in Smith 2003: 244). Moreover, such an interpretation is backed up by modal *must* in the previous sequence and it conveys the

utterance with strong force. Although preceded by first person pronoun subject *I*, the necessity refers to a strong speaker-involved comment on the specific situation.

6.3.1.4 Need to

The subjective-objective dichotomy with *need to* in present-day English has been given little attention in the literature, so far (e.g. Nokkonen 2006; Smith 2003; Taeymans 2004; Westney 1995). Despite the slightly different semantics of *need to* as compared to *must*, *have to* and *(have) got to* the distinction according to the source of necessity is useful for this marker, too (see section 4.5.3). Nevertheless, the classification according to discourse-internal vs. discourse-external sources requires minor modifications (cf. discussion in Nokkonen 2006: 62-64).

According to previous research, the semi-modal *need to* reports on a necessity which is internal to the speaker/subject "as an indirect marker of external necessity" (Taeymans 2004: 113). Such a source is illustrated in (109):

(109) Anyway I **need to** go up (ICE-Ireland.S1A-067.txt)

This example illustrates the basic meaning of *need to* where the necessity resembles an inner urge to the subject but which is determined by external factors (van der Auwera & Plungian 1998: 80). Examples (110) and (111) are similar however displaying both overt external and internal private obligations (Smith 2003: 245):

(110) And to finish up my M Phil in linguistics I have exams next week <#>Wish me luck <#>My <{8><[8>gosh</[8> I **need to** go study (ICE-JA.S1A-020.txt)

(111) I can't find that anywhere and I **need to** get it for my course (ICE-Ireland.S1A-016.txt)

Both examples report on a necessity which, at first reading, seems to be speaker-imposed. In spite of the speaker's internal motivation indicated with first person subject pronoun *I* as the source of necessity to study, the semantic reading is to be analysed in relation with a particular circumstance, i.e. exams approaching soon, or an existing requirement.

In (112) and (113) the speaker identifies herself/himself with the subject which is again in the first person:

(112) I don't want to be well known<{5><[5><,></[5> but you know to have my children I **need to** be wealthy (ICE-JA.S1A-026.txt)

(113) The only area that we **need to** have some improvement in is tourism and if the country will just place a greater<,> interest in eco-tourism we will also do very well in that (ICE-JA.S1A-049.txt)

As Nokkonen (2006: 46-47) observes, such uses do not refer to self-compulsion but emerge as the speaker's needs in the context of specific external conditions as shown above in (108) and (109). Example (110) illustrates such use in a rather ambiguous way as it is not very clear to what extent the necessity denotes the speaker's internal desire or merely volition. The second interpretation in fact brings *need to* closer to the semantics of *must* and even *(have) got to*. It is such isolated usage that, predictably, enhances the idiosyncrasies in New Englishes.

By contrast, in (113) first person plural *we* is rather a rhetoric device which underlines the speaker's comment on an existing necessity. Since the subject cannot be identified with the authority in charge to fulfil the necessity, such utterances are weak and objective. Furthermore, there are cases when first person subjects in the plural appear to identify with the source of the necessity, however, most often *need to* entails an instruction as found in (114):

(114) Now we **need to** enter what the constraints<,,> (ICE-JA.S1B-007.txt)

Although the speaker seems to address a directive, here the use of *we* in correlation with *need to* is impersonal as it reports on an external necessity related to an instruction for solving a problem in class. By contrast, in (115) the speaker reports the conditions of complying with a circumstance to another group (cf. Nokkonen 2006: 47):

(115) We have some gaps in some of the categories and we **need to** fill out these gaps and apply it across the board to other people but we are willing to look at these discreet cases (ICE-JA.S1B-073.txt)

The reason for using *need to* here seems obvious. As this example is extracted from the text category 'business transactions', the semi-modal is a more polite option to convey

objective necessity in a situation where the parties involved wish to reach an agreement by considering the interests of both sides.

Analogous with the case above, in (116) it is second person *you* which is impersonal:

(116) Now of course you **need to** find out what kind of a recording you can get with a tape recorder of whatever size and whether you need a mike or whether you don't need a mike ICE-Ireland.S1B-001.txt)

The speaker, most likely the lecturer, uses second person *you* to draw the attention of his audience to specific conditions which might influence the recording of sociolinguistic interviews. Although the speaker id here the expert, the source of necessity expressed with *need to* is not imposed and therefore is determined by objective factors.

The objective element of *need to* in (117) appears to be quite untypicaly, in which the speaker uses as agent first person *I* as a rhetoric strategy to emphasise the importance of the applicability of specific technical procedures:

(117) Alright now <}><->the</-> <=>the</=></}> vocabulary is talking but what I **need to** create is the ability to modify new strings (ICE-JA.S1B-009.txt)

The necessity seems, therefore, linked both to a condition and to "the internal logic of the discussion" (Westney 1995: 112). Since the context of such utterance is a lecture, we may guess that the speaker's intention is educational and, therefore, the necessity lacks the force of a concrete demand. Such uses are devoid of the basic meaning of the semi-modal thus pointing to increased abstraction (see discussion in Leech et al. 2009: 110-111).

Next, example (118) exploits a hypothetical situation as related to an internal necessity of the semi-modal originating in the physical needs of the human body:

(118) Ah God no <#> It would be dark <#> It would look romantic <#> I **need to** go to the loo (ICE-Ireland.S1A-050.txt)

In his discussion, Westney (1995: 107-108) points to such uses particularly with *must, have to* and *(have) got to* but omits mentioning *need to*. Although semi-modal *need to* "indicates a compulsion which comes from *within*" (emphasis as original) (Perkins 1983: 62-63), the discourse-external interpretation of the utterance above is related to a

need which the speaker has no conscious control over. The most appropriate strategy to test the internal motivation of such a reading as lying outside the speaker's control is to negate it (e.g. I *do***n't need to** go to the loo), obtaining lack of necessity related to bodily functions (Perkins 1983: 63). It is particularly such contexts which reveal the main distinction between *need to* and the other three markers under scrutiny.

Apart from the characteristic inner urge meaning, the objective interpretation can be related to a condition. Thus, all examples from (119) to (121) are compatible with the paraphrases 'X is necessary *in order to* Y' and '*if Y is to actualize*, X is necessary' (Depraetere & Verhulst 2008: 8; Wärnsby 2009: 91):

(119) Because some <,> students come into the schools <{1> <[1> knowing a language </[1> which is not English <#> </{2> <[2> And </[2> they **need to** understand it as a second language <{3> <[3> so as </[3> to be able to learn it in school (ICE-JA.S1A-001.txt) [i.e. *'if one wants to learn English in school'*]

(120) Or in other words <,> uh if you don't know English <,> if you want to communicate with people from other cities you **need to** know around uh ten fifteen languages <,> (ICE-India.S1A-025.txt) [*if*-clause, conditional]

(121) So that 's why you **need to** get there early and get get a decent seat (ICE-Ireland.S1B-016.txt) [i.e. '*if you want a decent seat'*]

In these examples the conditioning factor arises out of the actualization of a necessity: to understand English as a second language in (119); to know several languages in (120); and to come early to class in (121). A further aspect worth mentioning is that there is lack in temporal reference between the conditioning factor (the protasis) and the actualization of the requirement with *need to* (the apodosis). Incidentally, Westney (1995: 105) exemplifies non-specific reference with *have to* in contrast to *must* and *(have) got to* arguing that only the former implies present-future contrast. However, judging from example (121) such a distinction applies also to semi-modal *need to*, i.e. *you'**ll need to** get there early*. All in all, it is the pragmatic factors which strengthen the force of *need to* in spontaneous interaction.

6.3.2. Discourse-internal sources: speaker-imposition vs. personal opinion

6.3.2.1 Must

The discussion of root realization with modal *must* continues with the occurrences where the speaker is considered the source of necessity. For example, the source of necessity in instance (1) extracted from ICE-JA was described as originating from the speaker, which is at the same time the target, here first person subject pronoun *I* (cf. Westney 1995: 104). Such an utterance can be paraphrased as "*X* [i.e. to have an idea about the essay when one reads the thesis statement] *is necessary and my reason for saying so is that I am convinced it is necessary"* (Depraetere & Verhulst 2008: 11). Examples (122) and (123) are akin, as the source of necessity is internal to the speaker, who urges the addressee to perform an action:

(122) Uh it can prosper <,,> uh otherwise uh our country is <,> uh you know that it is going through a very <,> uh critical <,> time <,> and uh in order to survive that we **must** join our our hands together <,,> (ICE-India.S1A-005.txt)

(123) Obviously on the European circuit<{2><[2><,></[2>uh<}><->we</-> <=>we</=></}> **must** have some efficient coaches (ICE-JA.S1B-026.txt)

Although it is the speaker who imposes the necessity, first person *we* suggests identification of the source with the target. From a performative perspective, the uses of strong root meanings, where the speaker is identified as the source but does not identify with the addressee, are considered as "psychologically prototypical" (Coates 1983: 38). Examples (124) and (125) are relevant instantiations for such root meanings:

(124) Remember you **must** have some theory guiding your research also right right<,><O>students-commenting</O> <#> (ICE-JA.S1B-010.txt)

(125) You really want to get into the subject of what bioterrorism is right <#>So first you **must** have a good working definition of bioterrorism <{1><[1>right</[1> <#> (ICE-JA.S1B-005.txt)

In addition, the subjective element may be linked to a 'personal opinion' (Depraetere & Verhulst 2008: 12-14) where the necessity is typically conveyed by means of hedges (*I think, I believe, in my opinion* etc.) as in (126) to (133):

(126) And it's a bit of a paradigm shift because *we think* that our university **must** have four walls or a wall <?>insert</?><{15><[15><,></[15> you know going around it and so on (ICE-JA.S1A-019.txt)

(127) *She feel* she **must** be the hottest thing so <{5><[5>mummy</[5> did <indig>haffi</indig> carry her go hairdresser (ICE-JA.S1A-026.txt)

(128) That is no harm in growing the crops <,> <{4> <[4> uhm <,> but these things **must** be in a right order *I feel* <,> *I feel* like this <,> <{5> <[5> uhm uhm <,> uh <,> <{6> <[6> okay <,> because all these three must three things must come in a <,,> right line <,> then only you I feel that <,,> we can grow agriculture (ICE-India.S1A-084.txt)

(129) But of course Eleanor <,> uhm I I yes *I think* I I **must** admire Eleanor because uhm uhm <,> she has <&> laughter </&> <.> sh </.> she has immense <,> skill and talent (ICE-Ireland.S1A-061.txt)

(130) But *I think* <}><->as a</-> <=>as a</=></}> people we **must** do better than that (ICE-JA.S1B-050.txt)

(131) You know *I believe* that the union budget for nineteen ninety-one ninety-two <,,> **must** be situated in the wider context <,> of the crisis in the economy <,> that is both acute <,> and de (ICE-India.S1B-023.txt)

(132) But <,> other way *I think* is we **must** do something for <,> the uplift of the people <,> (ICE-India.S1B-025.txt)

(133) We're here representing the views of the people and we will not change our minds but *we believe* that the guns **must** go for the future of our country (ICE-Ireland.S1B-055.txt)

In spite of the hedges, which arguably suggest a softening of the strong compelling force related to modal *must*, these examples denote urgency of obligations through direct speaker involvement, without having the authority to impose them (Depraetere & Verhulst 2008: 13). Exceptions refer to a speaker which merely acknowledges the necessity of a university to have four walls as in (126), the impression someone has about her own person in (127) or the internal disposition which motivates the speaker to admire a particular person as in (129). The fact that modal *must* is compatible with the expression of an endorsement suggests pragmatic weakening. At the same time, the subjectified meaning emerges from the conjecture of the expert knowledge of the speaker (Goossens 2000: 164) as in (128), and in (130) to (133). On the other hand, the combination with first person *we* might indicate an agreement or consensus on the

necessity of an action involving shared interests, and is similar to Giltrow's (2010: 45) findings on the function of this modal – though without hedges – in academic writing.

A particular case is example (127) extracted from the Jamaican informal dialogues where the speaker is the channel of necessity conveying self-compulsion to the utterance. What is more remarkable here is the co-occurrence with *haffi*, as an orthographic variant of creole *hafi,* which in this case appears to add emphasis to the utterance. Judging from such alternation, it could be assumed that both the creole form, which arguably stands for standard *have to* (see also discussion in 4.5.3), and modal *must* convey the same meaning of obligation in educated standard Jamaican English. Such an interpretation then does not appear to warrant the overall high frequency of occurrence with semi-modal *have to* in ICE-JA. However, since this is an isolated occurrence it cannot be concluded that the two forms report the same semantic reading. Furthermore, in absence of related examples such interpretation is bound to remain tentative.

The next examples are in stark contrast with the assumed compelling obligation if a second person pronoun precedes the modal as in (134)[98] to (136):

(134) you **must** let me photograph your baby for my magazine (ICE-GB.S1A-039.txt)

(135) Yes certainly at least you **must** uh <,> attend the <,> club once (ICE-India.S1A-001.txt)

(136) You **must** take us </[> </{> to Banglore someday introduce us to your husband <{> <[> as well as </[> your son (ICE-India.S1A-031.txt)

Although the use of *must* in these three examples is compatible with a strong directive, the pragmatic context denotes a meaning related to wishes or invitations for the addressee to perform an action. While such overt insistence more resembles an imperative, these represent a social convention in present-day English (Palmer 1990 [1979]: 73; Westney 1995: 117). In view of such an interpretation the two examples from ICE-India converge with the standard English norm as illustrated in (134) in ICE-GB.

Equally strong is the obligation when the subject is in the first person, such as in (137) and (138):

[98] For alternative interpretations of the modal reading of this utterance, see Collins (2009a: 35) and Depraetere & Verhulst (2008: 12).

(137) I must ring Kevin (ICE-Ireland.S1A-042.txt)

(138) <S1A-050$A> <#> <[> You could </[> </{> <unclear> </unclear> me a hat then<S1A-050$C> <#> I will <#> I must knit you a hat anyway
<S1A-050$A> <#> You must do (ICE-Ireland.S1A_050.txt)

In these examples it is the speaker who imposes the necessity on herself/himself, which, again, is similar to examples (82) and (93), even though external factors prevail over the internal disposition in those two cases. Apart from strong obligation, such uses appear to be compatible with an element of volition (Jespersen 1924: 320). In (137) it is a spontaneous decision which conveys the utterance self-compulsion (Westney 1995: 104), whereas in (138) the alternation with *will* suggests future reference of the obligation with the addressee in the following speech turn confirming the speaker's intention through an invited inference. Overall, discourse-internal sources with modal *must* increase the subjectivity involved in the obligation.

6.3.2.2 Have to

If discourse-internal sources are usually related to a directive, in the following I shall exemplify on the basis of corpus data that the imposition of obligations and necessities as identified with the speaker may occur also with the semi-modal *have to*. The examples below selected from both formal and informal dialogues from ICE-JA, ICE-India and ICE-Ireland suggest that in the absence of a discourse-external source a subjective interpretation is most appropriate. It appears that, as with *must*, most often the subjective interpretation is conveyed by the speaker's implication as an opinion or recommendation and thus confirms the findings from Depraetere & Verhulst (2008). In addition, the present corpus data from New Englishes shows that clear speaker involvement as a directive is also possible with this semi-modal. Utterances (139) to (161) stand as evidence for such root realizations in JamE, IndE and IrE. A more detailed evaluation of these examples might suggest that prototypical root meaning with semi-modal *have to* has extended to become more specific in New Englishes:

(139) In uhm </-> <=> on </=> </}> the side of acknowledging this I mean </{4> <[4> that </[4> is we **have to** develop this thing <{5> <[5> <unclear> a-few-words </unclear> </[5> as part of the culture <#> (ICE-JA.S1A-001.txt)

(140) They **have to** bathe then come in<O>laughter</O> <$?><#> (ICE-JA.S1A-011.txt)

(141) And actually the first year <}><->I</-> <=>I</=></}> applied for UWI I applied for management studies<{5><[5><,></[5> and didn't get in <#>So the second year I said <#>Now I **have to** get in <#> (ICE-JA.S1A-027.txt)

(142) So I hope I'll get to see it because I really Tereka I really love singing any show that has to do with singing trust me no matter what I'm doing no matter where I am no matter who I am with I **have to** see the singing especially when it's good singing and good music <#> (ICE-JA.S1A-033.txt)

(143) But <}><-><.>wha</.></-> <=>we</=></}> **have to** accept that you know because<,> at some point in our lives good friends have to part and it's something that we all have to get used to because no matter where you are somebody you get close with at some point in time they either migrate or gone somewhere else to work or (ICE-JA.S1A-033.txt)

(144) Uhm come back to the point *I think* as a country who have come out of the English uhm tradition we **have to** decide now is this<,> the first language and English a second language<{8><[8><,></[8> (ICE-JA.S1A-039.txt)

(145) <[> No actually no </[> </{> there **has to** be a communication with the students and the teacher <,> I *I find* that I have a gap <,> with my students <,> (ICE-India.S1A-085.txt)

(146) UWI is so theory-based when you leave here <}><->you</-> <=>you</=></}> definitely **have to** do your masters<,> to go<?>practice/back to</?> some things <{18><[18>or</[18> go to a company that is definitely teach you what they want you do < (ICE-JA.S1A-047.txt)

(147) Uh J N U in the evening <,,> uh you **have to** see the J N U <,,> uh you can visit until seven <,,> people thronging the bus <,> and they talk Marxism <,,> art <,> cultural dance debating <,> so all those uh <,> smart<,,> girls and boys in J N U <,,> used to sit together and they used to talkas <,> not all <,> not all of them <,,> (ICE-India.S1A-028.txt)

(148) Perhaps the time has come <,> when *I think* across the board <,> we **have to** take a view <,> whether it is Kashmir <,> or Andhra <,> or Assam or <O> one word </O> <,> or anywhere else <,> (ICE-India.S1B-024.txt)

(149) <S1A-025$B> <[> Nora </[> </{> who 're you going to marry
<S1A-025$A> <#> I don't know <#> **Have to** see <#> Won't be marrying anybody sexist anyway (ICE-Ireland.S1A-025.txt)

(150) I just **have to** get into a routine <,> a regular routine <#> And I I know if I kept up that type of pattern I 'd be fine <#> It 's just starting it (ICE-Ireland.S1A-025.txt)

(151) Eleanor you **have to** be cruel to be it 's <.> cr </.> you *have to* be cruel to be kind in the right measure <&> laughter </&> (ICE-ireland.S1A-061.txt)

(152) You **have to** play Simon Says <#> (ICE-Ireland.S1A-088.txt)

(153) You **have to** make sure that's what Clarisse wants as well (ICE-Ireland.S1A-092.txt)

From these examples it emerges that subjective meanings with *have to* are most subtle in contexts of personal opinions. It is the speaker who expresses their beliefs or opinions about the necessity of the country to distinguish between first and second language in a traditional British influenced country as in (144); to communication between students and teacher as in (145); and to take a wider perspective in problems related to Indian politics as in (148).

Intuitively, the fact that the semi-modal *have to* is usually associated with a general requirement lying outside the disposition of the speaker makes it a more appropriate choice in such contexts. In other words, stating a necessity with *have to* which reflects the speaker's personal opinions or beliefs is at the same time a useful strategy to shift the attention from the speaker to the actualization of necessity itself. At the same time, it seems to be more appropriate than the modal *must*. While both items yield equally strong obligation, it is the hedges in combination with the semi-modal in this case which increase the speaker's commitment to a situation (see discussion in Depraetere & Verhulst 2008: 13). As I suggested in section 4.6.3, such uses can be interpreted as representing an intermediate stage towards even more subjective uses thus involving the conventionalization of implicature. Taking into account the importance of subjectivity in the development of modal *must* it might seem that the semi-modal *have to* follows a similar path of semantic shift (see more detailed explanation in section 6.3.3.2).

Notable, the semi-modal *have to* is not restricted to only expressing an expert opinion. Proof of speaker-imposed obligation are found in examples (139), (140), (141), (142), (143), (149), (150). Even among these however, can uses where *have to* refers to a self-imposed obligation having narrow scope be identified, like in (141), (142),

(149),[99] and (150) as compared to a directive in (139), (140), (143) – admittedly, these instances denote rather neutral obligation – and (146). While semi-modal *have to* can be used in directives it seems a less common choice for self-compulsion. Instance (141) increases speaker involvement by displaying the same strong effect as modal *must* in the sense of a specific requirement with the speaker urging herself/himself. An appropriate paraphrase of this utterance is then *'it is obligatory/absolutely essential for…'*. Notable in this case is that such use of *have to* seems to involve volition also observed for modal *must* (see section 6.3.2.1). Conversely, the semi-modal in (150/157) underlines the habitual aspect of *a regular routine*, which involves a self-imposed necessity.

A particular instance of pragmatic weakening with *have to* is (151). Even though the speaker is addressing a directive, it seems to be more compatible with a recommendation. From this perspective, the semi-modal is closer to the semantics reported with *need to*. The corpus data shows that such uses with semi-modal *have to* are equally preferred in formal and informal contexts of speech. In addition, as with modal *must* in (134) to (136), the semi-modal can express a polite way of addressing an invitation such as (152).

Furthermore, similar to *ought to* in Nordlinger & Traugott (1997), in example (145) the subjective element imposing an obligation with *have to* entails a wide scope interpretation of the modalized proposition. Despite the use of hedges signalling personal implication, the obligation in this example emerges out of specific expectations. In addition, the obligation is supported by an unspecified agent. As with the interpretation of the path from root/deontic to epistemic modality with other modal items (cf. Nordlinger & Traugott 1997: 304), the various degrees of subjectivity in the uses of root *have to* in synchronic data from New Englishes can be approached in terms of scope as a further parameter of ongoing change.

To sum up, as emphasised in Depraetere & Verhulst (2008: 14), the present data suggests that subjectivity with *have to* is more compatible with the expression of a personal opinion or a recommendation which "is necessary for a certain situation to be brought about without actually giving a directive." In this sense, it is most likely that the

[99] In spite of the elements justifying an interpretation of self-imposed obligation, such use of *have to* contrasts with *must* in that it might equally function as a strategy of the speaker to distract attention from himself/herself in conversation.

increase of such uses in the present data is related to a process of democratization (see section 4.5.2).

6.3.2.3 (Have) got to

In section 6.3.1.3 I emphasised the near equivalence between *(have) got to* and *have to*. Excerpt (153) extracted from the private dialogues of ICE-India illustrates such an alternation between these two items:

(153) And uh they **got to** go to <,> they *have to* go to different industries <,,> where <,,> the students are expected <,> to go <,> and do some industrial training (ICE-India.S1A-023.txt)

Given the low frequency of this semi-modal in the Indian sub-corpora in general, such alternation is most likely to be interpreted as a false start. On the other hand, it might also be a case of conscious self-correction. While both *got to* and *have to* are originally informal and colloquial, it is remarkable that the speaker opts for the latter form, which is also the established form in present-day English at the expense of the most innovative one.

As the examples below show, subjective *(have) got to* exhibits somewhat different semantics than *have to*, as it expresses a more specific requirement and adds urgency or conveys emphatic affirmation to the utterance (Westney 1995: 112). Thus, in contrast with modal *must*, *(have) got to* expresses "obligation motivated by emotion" (Myhill 1995: 163). Such motivation is most compatible with a subjective/discourse-internal source with speaker imposition as in example (154):

(154) See <{> <[> we **have** </[> **got to** take the responsibility of the fault (ICE-India.S1A-018.txt)

In accordance with the subjective interpretation, instance (154) suggests speaker involvement in the form of a directive addressed to an audience, in this case the source identifies with the addressee using first person plural *we*. The urgency of the situation is very clear in this utterance with the speaker urging the addressee to undertake a specific action, i.e. *the responsibility of the fault*. The use of *have got to* in this utterance might

be tentatively interpreted as the speaker's appeal to the addressee's moral conscience (cf. also the examples from (160) to (166)).

By contrast, the subjective element in (155) and (156) converts itself in self-imposition of necessity as was already shown both with *must* and *have to*:

(155) so I thought right <#>You '**ve** <{1> <[1> **got to** try </[1> this (ICE-Ireland.S1A-044.txt)

(156) Well I **'ve got to** have my hair nice for tonight (ICE-Ireland.S1A-059.txt)

However, some minor differences emerge from these two examples. While the necessity in (155) refers to a specific requirement the speaker addresses to a hearer, and which could be equally stated with *have to*, instance (156) suggests that the necessity originates in the speaker's internal disposition to face a situation, which again enhances urgency. On the other hand, second person pronoun subject *you* in (155) can be seen as a discourse strategy that is often used in self-encouragement and is along the same lines as (156), however with the speaker, i.e. first person *I*, as the source urging her/himself to comply with the necessity.

If *(have) got to* can express self-compulsion both with a subject in the first and second person as shown above, (157) addresses a strong requirement to the audience to remind the speaker about an important event:

(157) Lads you **'ve got to** remind me to meet Emma at five o'clock next Tuesday (ICE-Ireland.S1A-066.txt)

Although speaker involvement is overt in this example, the utterance does not impose any obligation on the addressee. Instead, such uses emphasise the speaker's commitment to the necessity.

In the same tone, *(have) got to* can express friendly and polite inquiry. Thus, the use of this semi-modal in (158) is semantically close to the subjective motivation with *must*:

(158) Monica I **'ve got to** get your phone <,> or your mobile number (ICE-Ireland.S1A-019.txt)

Such uses typically occurring with modal *must* suggest that they have extended also to other variant forms in the system.

According to Westney (1995: 110), particularly the combination with the first person subject pronoun *we* displays "the strongest type of urging, more directly than with *must* and *have to*", and notes that in British English it is most typical in argumentative discourse in public announcements. A similar occurrence is found in utterance (159) selected from a broadcast discussion in ICE-JA:

(159) This society **has got to** come <{1><[1>to terms</[1> with itself (ICE-JA.S1B-029.txt)

The motivation behind the necessity in this example is related to a strong requirement. Although modal *must* would underline the strong moral principles related to such a necessity, *has got to* in correlation with inanimate definite subject *this society* emphasises the urgency of the speaker-imposed requirement.

Finally, the present data shows that subjective *(have) got to* is compatible with the expression of personal opinions particularly in public statements. Significantly, such uses typically occur in public contexts of interaction such as in 'broadcast discussions' in (160) and in 'broadcast interviews' in (161):

(160) *I think* that is the point <w> we'**ve** </w> **got to** understand (ICE-India.S1B-033.txt)

(161) Again time and again we have dealt with this problem and <,> <{> <[> uhm <,> and on radio <,> *I think* <w> you've </w> you **have got to** do it (ICE-India.S1B-025.txt)

In addition, the hedging element (*I think*), occurring in all the examples, has the same effect of a less threatening obligation as discussed earlier with modal *must* (see 6.3.2.1) and semi-modal *have to* (see 6.3.2.2). At the same time plural reference emphasises the necessity which the participants share in this situation. A similar search in ICE-GB suggests that such combinations are indeed reserved for public discourse such as 'broadcast discussions' in (162) and (163), and 'broadcast interviews' in (164) and (165), and even in 'business transactions' in (166):

(162) but equally you *I think* you **'ve got to** pick up the point about how efficient the council is (ICE-GB.S1B-034.txt)

(163) Nonetheless I mean *I think* the Government uh **have got to** show that they are able to produce and not just slag off the Labour Party (ICE-GB.S1B-039.txt)

(164) And uh from having come from A T C where I was working on classical texts by people like Molière and Shakespeare one of the other things *I feel* strongly about is that we **'ve got to** get writers back into the theatre more involved with actors (ICE-GB.S1B-050.txt)

(165) *I think* what you **'ve got to** do is you 've got to try and set up again for people a kind of of of a vision of theatre and uh a vision of creating theatre (ICE-GB.S1B-050.txt)

(166) however *I think* we **'ve got to** make more effort more more of an effort to actually encourage students to come and to uh fill the places if we 're going to in increase it (ICE-GB.S1B-075.txt)

From the data found in ICE-India, therefore, despite the low frequency of this semi-modal, IndE closely follows the British norm in public interaction.

6.3.2.4 Need to

The discussion of discourse-internal sources ends with a description of the uses of *need to*. While the basic meaning of need to is confined to the internal disposition of the speaker as determined by external factors (see 6.3.1.4), in this section I will explore the subjective uses of this marker in the New Englishes under study. An examination of such root readings is to be seen in the light of conventionalization of implicature through pragmatic inference in the ongoing development of this marker (cf. Nokkonen 2006: 67; Taeymans 2004: 108).

In contrast with *must*, *have to* or *(have) got to*, the subjectivity of need to is pragamtically inferable to which Smith (2003: 245) refers as "a potential for speakers to exploit this internal quality of NEED (TO) in order to obtain some advantage" (emphasis as original). The reading of such utterances would point to an extension of meaning towards increased speaker-based necessity. Several examples in the database display such overlap in meaning, suggesting increased subjectivity, at a more subtle level, however, as already exemplified in (44). As with the other markers, subjectivity with need to is related to subject selection, such as with second person you in (167) and (168):

(167) Uhm what you **need to** do restructure the thesis statement<{1><[1><,></[1> (ICE-JA.S1B-005.txt)

(168) You **need to** stay with me and do it so you know what I am telling you <unclear>word(s)</unclear> I'm going tell you to do (ICE-JA.S1B-005.txt)

In these examples second person *you* clearly indicates the existence of some authority, such as a teacher. More than in (167) where the speaker uses *need to* to address a recommendation/to give an advice about how to comply with a task, (168) conveys the utterance a speaker-imposed directive, which almost sounds like a command. Interestingly, such formulations are common for the text category of 'class lessons' in ICE-JA and ICE-Ireland.

Another possibility of expressing subjective motivation is with first person subject *we* which, on the one hand, suggests self-identification of the speaker with the audience, but on the other hand, hides a directive as in (169) and (170) (cf. also example (44)):

(169) We **need to** be looking at the schedule of benefits that are in relation to those jobs at present because there might be other things (ICE-JA.S1B-073.txt)

(170) No We **need to** pray for me I feel sick <,> I think I 've got bellyache after after </> Wonkie 's announcement (ICE-GB.S1A-070.txt)

As noticed elsewhere in Nokkonen (2006: 47), both instances refer to a strong directive reading among equal participants in interaction as, for example, in business transactions in (169) or with friends in (170). In addition, example (169) displays a further grammatical aspect worth mentioning: the fact that the speaker uses *need to* in combination with a progressive construction (i.e. ***need to*** *be looking*) adds significance by conveying commitment to the actualization of necessity.[100] A comparatively similar use with *have to* is found in (171) also selected from the public dialogues in ICE-JA:

(171) We are saying that this should come out of some other budget but we **have to** be dragging this out of you in this budget this should be fund money we should be negotiating for <$C> (ICE-JA.S1B-080.txt)

In contrast to (170), where the directive is expressed rather indirectly, the combination of *have to* + be-*ing* in (171) is closer to the interpretation of speaker commitment and

[100] Consider also the combination *will* + be-*ing* which implies that "the projected event is in keeping with what the speaker considers to be normal, and/or free of volition or intent" (Leech et al. 2009: 143). On the evolution of the progressive in combination with a modal auxiliary in present-day English, see Leech et al. (2009: 139-141).

thus of enhanced subjective obligation. Conversely, *need to* in (172) followed by stative *look* is still subjective, however the directive function is less strong:

(172) The second factor we **need to** <,> look at in this budget is the context in which it was written in a sense <,> (ICE-India.S1B-023.txt)

The subjectivity in utterances involving authoritative stance, such as personal directives or comments, suggest a strong involvement in combination with hedges (Collins 2009a: 73) as in (173) to (176), and as was also shown with the other markers:

(173) *I think* you **need to** conclude shortly Sine/ad (ICE-Ireland.S1B-020.txt)

(174) *I think* we need more practicality uhm we **need to** be out in the field more with certain things<{4><[4><,></[4> and I know that probably will cost more and stuff but we need a little bit more of that (ICE-JA.S1A-016.txt)

(175) Therefore <w> it's </w> not just Kashmir <,> and *I think* we **need to** look at this not as isolated problems (ICE-India.S1B-024.txt)

(176) But equally uh *I think* it 's implicit in the process though I I would have welcomed a much more explicit recognition by the British government <,> that the the Nationalists do have grievances that they do have a sense of identity that **needs to** be recognised and given some institutional form in whatever new arrangements are put in place even if they 're only interim arrangements moving towards a solution and none of us is is sure what the the the final outcome will be (ICE-Ireland.S1B-021.txt)

According to Nokkonen (2006: 49-51), a special use of *need to* arises in combination with passive constructions, which do not allow for a clear-cut interpretation of the source of necessity as the agent is not always overt (cf. van der Auwera & Plungian 1998: 101). While in her corpus, strong subjective uses occur mostly in fictional texts (Nokkonen 2006: 49), in the present data these occur in argumentative contexts from public dialogues as in (177) to (179):

(177) And public opinion ultimately must prevail in these matters <{> <[> and that needs </[> to be educated <$A> (ICE-India.S1B-024.txt)

(178) In the century to come <,> this isssue of <,,> population aging <,> will raise many vital and new questions which **need to** be addressed urgently <,> (ICE-India.S1B-035.txt)

(179) But the fact of the matter is it 's the system in which we operate which is totally archaic and it <{> <[> **needs to** be fundamentally changed </[> (ICE-Ireland.S1B-031.txt)

In all these examples the source is the speaker, who can be identified as the authority urging their audience as all the subjects (i.e. *public opinion* in (177), *new questions* in (178) and *the system* in (179)) correlating with *need to* are inanimate. In addition, in (177) *need to* alternates with modal *must*, thus adding more subjective force to the utterance. The impression given by this utterance is that the speaker is appealing to his/her expert knowledge to emphasise the needs of the audience. It follows from this that *need to* is a more polite option than the other markers. By contrast, the speakers in (178) and (179) add urgency to the actualization of the necessity very much in the sense of *(have) got to*, which is grammatically conveyed with the two adverbs, *urgently*, in the former, and *fundamentally*, in the latter. While the majority of such uses occur in the formal spoken section of the Jamaican, Indian and Irish corpora, in the first corpus it occurs also in informal contexts of interaction. As Loureiro-Porto (2009: 173) observes, such uses of *need to* are a reminder of modal *need* in combination with passive voice as found throughout history from OE to ME, and which syntactically and semantically resemble the modals. From the perspective of present-day English, the correlation with the passive is suggestive for the way in which *need to* selects the subject, which is similar to the modals. Likewise, the fact that these combinations occur more often in formal contexts of speech in New Englishes might be indicative of an ongoing development (Nokkonen 2006: 60 f.).

6.3.3 Disentangling variation: discussion of corpus findings

6.3.3.1 Regional and stylistic trends

Having described the semantic range of sources with these four markers in samples from three different regional settings, it becomes apparent how difficult it is to draw neat categorisations of usage patterns in natural language data. The aim of describing root necessity was to emphasise the contrast between subjective vs. objective modality and how such an approach can account for ongoing change in the three focus varieties,

JamE, IndE and IrE. However, evidence of possible shifts in the meaning of these modals has been mentioned only on the basis of individual instances. In the following, I shall summarize the main corpus findings pointing to the principal patterns for root meanings with *must*, *have to*, *(have) got to* and *need to* in each variety.

Overall, the description of sources as discourse-internal or discourse-external has proved to be a useful heuristic tool. In hindsight of the initial classification, the previous assumptions on the subjective vs. objective distinction (see Fig. 6.6) call for minor adjustments. The most significant finding is that root necessity covers a wide range of semantic distinctions which can be summarised roughly as in Table 6.1:

Discourse-external	***Discourse-internal***
internal desire/wishes	(hidden) directive
rules/regulations	recommendation
circumstances/conditions	personal opinion
	(expert knowledge of the speaker)

Table 6.1. Sources of root necessity in the present data

The table above (Table 6.1) corresponds in fact to the previous classification in section 4.5.3 to which further uses were added, e.g. internal desire/wishes which mark external motivation and directives as originating in the speaker. As the data shows, such meanings are not restricted to *need to* as the prototype item expressing internal disposition of the agent as determined by external factors, but under specific circumstances are conveyed by the other three markers as well. Such a finding emphasises the potential overlap of variant forms to express root necessity. At the same time, these sources illustrate the steps involved in the process of subjectification which consist of "a multiplicity of paths conspiring to establish a new conventionalized (prototypical) use" (Goossens 2000: 167).

More than the identification of sources though, it is the regional distribution of these contrasts with root necessity that are of interest here. For reasons already mentioned, the present assessment includes information only from ICE-GB as representing contemporary BrE. In addition, the distinction according to text categories and level of formality in each of the corpora under study provide the relevant dimensions to assess such processes in interaction.

To begin with, Table 6.2 and 6.3 cover the distribution of sources with all the four markers jointly in formal and informal contexts of interaction in ICE-JA, ICE-India, ICE-Ireland and ICE-GB:[101]

	Discourse-external %			
	ICE-JA	**ICE-India**	**ICE-Ireland**	**ICE-GB**
must	17.1	23.5	8.7	11.2
have to	63.5	73.1	72.3	34.5
(have) got to	0.7	1.5	11.4	39.4
need to	18.1	1.7	7.6	14.6

Table 6.2. Distribution of discourse-external sources of root necessity in percent

	Discourse-internal %			
	ICE-JA	**ICE-India**	**ICE-Ireland**	**ICE-GB**
must	16.4	33.9	21.6	16.6
have to	55.2	57.4	51.3	41.3
(have) got to	1.5	1.9	9.5	24.1
need to	26.8	6.2	17.4	17.7

Table 6.3. Distribution of discourse-internal sources of root necessity in percent

The figures in the two tables account for the proportion in percent from the total amount of discourse-external and discourse-internal sources in each variety (see Appendix 2.c). As I pointed out elsewhere (see section 4.7.5), due to indeterminacy in modal meaning a clear-cut interpretation of modal necessity is often difficult to obtain. The same observation applies to the cases where objective sources report, for example, on circumstances or conditions motivating the necessity and which in some instances are not easy to distinguish, but where an external factor is still identifiable. Therefore, I do not provide separate counts for each type of sources but will discuss these under the general term of discourse-external sources.[102]

It must be also emphasised that, despite the regional differences in terms of raw frequencies, the distribution of sources within the total amount of root necessity with each of these markers (see Appendix 2.a) reveals rather balanced usages. Moreover,

[101] Again, a direct comparison of *must* and *have to* with the findings from Depraetere & Verhulst (2008) is rather limited, as the present evaluation is based on data from private and public dialogues in ICE-GB. Note also that a distinct category of 'merger' was not included here.

[102] For a detailed discussion on the evidence of such sources in ICE-GB, see Depraetere & Verhulst (2008: 19-23).

within such a configuration it is in ICE-Ireland and ICE-GB where associations of subjective necessity with *have to* are more pronounced than in ICE-Jamaica and ICE-India, where such uses show a preference for modal *must*. Although such a finding would lend support to most of the intuitions from the literature about modal necessity in the inner and outer circle (see Collins 2009b), it will be useful to recall that the total amount of root readings with individual items in each variety varies to a large extent (see section 6.2) and, therefore, a comparison yields inaccurate assumptions about the spread of such uses in these New Englishes. This is more problematic given the lack of other comparable attempts to classify subjectivity and objectivity of root necessity in JamE, IndE and IrE. Clearly, the frequency counts which are taken as evidence for the interpretation of the semantic behaviour of these items are provisional and do not allow for more advanced uses of root necessity in either outer or inner circle, and least so for language change.[103] Irrespective of this incongruity, the present section deals instead with the layering of root necessity with these items within each regional dataset. These observations will be relevant in identifying the general orientation in the expression of root necessity in the New Englishes, as compared to the established standard norm, i.e. BrE. Note, however, that such an approach does not provide sufficient reasons to sustain significance of variation in the distribution of subjectivity and objectivity in the inner and outer circle varieties either. Nevertheless, the figures above do not preclude the pointing out of some synchronic tendencies in the New Englishes.

Some interesting aspects emerge if when examining the sources as linked to one of the four items in each variety. Even though, overall, there is no clear association of these with a specific source, there are some differences between inner and outer circle varieties. When this distinction is viewed independently in the different datasets it appears that the beneficiaries of the increase in subjectivity are the modal *must* in ICE-India, the semi-modals *have to* and *need to* in ICE-JA, and *(have) got to* in ICE-GB. Moreover, provided the analysis has captured the subjective elements correctly, it emerges that such contexts are reported predominantly with *have to* which outnumber those with *must* in all the regional datasets. That this semi-modal is associated with discourse-internal sources in the dialogues from ICE-GB is also striking, but in view of

[103] Such discrepancies in the data are particularly challenging for a corpus-based study of synchronic variation if one assumes that ongoing change can be inferred from evidence in the apparent time. See also Baker (2010: 81-85) on reporting results in corpus-based studies based on synchronic variation.

the register under study is compatible with the findings from Depraetere & Verhulst (2008: 22).

As for objective meanings, ICE-India leads in regard to frequency of discourse-external sources with *must* and *have to*, ICE-JA with *need to* and, most interesting, ICE-GB with *(have) got to*. While ICE-JA displays larger proportions of subjective *have to* and *need to*, ICE-India and ICE-Ireland reveal a contradictory pattern, as both varieties appear to display a competition between *must* and *have to* in the use of discourse-internal sources. Except for ICE-JA, subjective uses with modal *must* outnumber objective uses in all the other datasets. This might indicate that subjective necessity in JamE has been taken over by other markers, such as *have to* and *need to*. On the other hand, both *must* and *have to* are associated more often with discourse-external sources. Finally, the fact that modal *must* is associated most often with subjective necessity in ICE-India might be linked to the high incidence of epistemic meanings in this dataset (see Figure 6.1).

Further contextual factors such as level of formality and text categories will help identifying the extent of layering in these varieties. Following this line of reasoning, the functional distribution in specific environments might be useful to account for variation at a specific point in time within each variety. Thus, both discourse-internal and discourse-external sources with *must* are preferred in formal contexts of speech in ICE-India with 71.4% and 63.3%, respectively, and ICE-Ireland with 83.3% and 60.9%, respectively. It is in the informal dialogues from ICE-JA and ICE-GB where discourse-internal sources are reported with the modal. The specific uses of *must* in formal contexts in ICE-India are noteworthy. In the description of objective sources as illustrated in the examples (72) and (73) I pointed to the combination of a generic/impersonal subject and an external force which conveys objective root necessity with this modal particularly in ICE-India, which is consistent with the text category of ‘class lessons’ (with 8 occurrences). Contrastively, in ICE-JA these uses occur rather in the ‘face-to-face conversations’ with at least six occurrences that match those in the IndE corpus mentioned above. The assessment of similar uses in ICE-Ireland and ICE-GB has yielded no occurrence. What is of relevance in this context is that, despite being a minor feature in the two corpora, such external motivations have contributed to the development of obligation meanings of *must* in the English modality system. In the

absence of similar evidence in the two inner circle varieties, the few occurrences in the JamE and IndE datasets might tentatively point to traces of older uses. On the other hand, this finding most likely indicates different cultural dimensions of necessities which are linked to an authority as represented by higher forces.

In regards to other text categories, further evidence for objective necessity is found particularly in 'broadcast discussions', 'broadcast interviews' and 'parliamentary debates', which are relatively evenly distributed also in ICE-GB. Interestingly, it is in the 'parliamentary debates' from ICE-India where the largest amount (22 occurrences) of such uses is found, and which generally report on rules and regulations. Furthermore, both ICE-JA and ICE-India show consistent use of objective necessity in informal dialogues. Likewise, the distribution of objective sources in formal dialogues covers all text categories in these two corpora. It is only in the 'business transactions' from ICE-JA where modal *must* expresses discourse-external sources related to institutional regulations but also to principles, whereas the speaker as source of necessity with this item in the same text category is less often used overall.

A further remarkable observation is that discourse-internal sources with the modal in informal dialogues in ICE-India and ICE-JA hardly occur performatively in the sense of prototypical obligation meaning (cf. Coates 1983: 33). Instead, most uses are correlated with first person definite subject *I* but also impersonal *we* and report on internal dispositions of the speaker in the sense of self-exhortation (see examples (1) and (138)) or a conventionalised use, such as an invitation (see examples (134) to (136)). This uncovers similarities with the British pattern found in ICE-GB and which is found also in ICE-Ireland. In addition, I found three occurrences in the dialogues from ICE-JA which correlate with a third person subject, usually *she*, and which indicate that the speaker is only the channel reporting the necessity (see example (127)). Overall, subjective uses in public speech are preferred in ICE-India with 63.3%. Furthermore, many subjective sources in the 'broadcast interviews' (four occurrences) and 'broadcast discussions' (two occurrences) from ICE-India are compatible with an interpretation of pragmatic weakening of the obligation meaning but which at the same time indicate the speaker's identification with the necessity such as in combination with a hedge as illustrated in the examples (128), (132) and (133). There was only one such example in the same text category from ICE-Ireland but none in ICE-JA. It is ICE-GB which

provides several examples of such forms spread over several text-categories, such as in 'class lessons' (one occurrence), 'broadcast interviews' (three occurrences), 'broadcast discussions' (two occurrences), 'parliamentary debates' (two occurrences) and 'business transactions' (one occurrence). Such combinations most likely point towards the formal use of this item. That subjective necessity in combination with hedges is productively expressed with *have to* in ICE-JA (109 occurrences representing roughly 62% of the total of subjective sources with the semi-modal) reveals an interesting finding in educated standard JamE and requires further investigation (see Figure 6.7). The fact that their use occurs predominantly in the formal text categories is suggestive of a possible democratisation in language in the way speakers exert authority on an addressee or express expert knowledge, as well as commitment to the necessity of an action (see section 6.3.2.2).

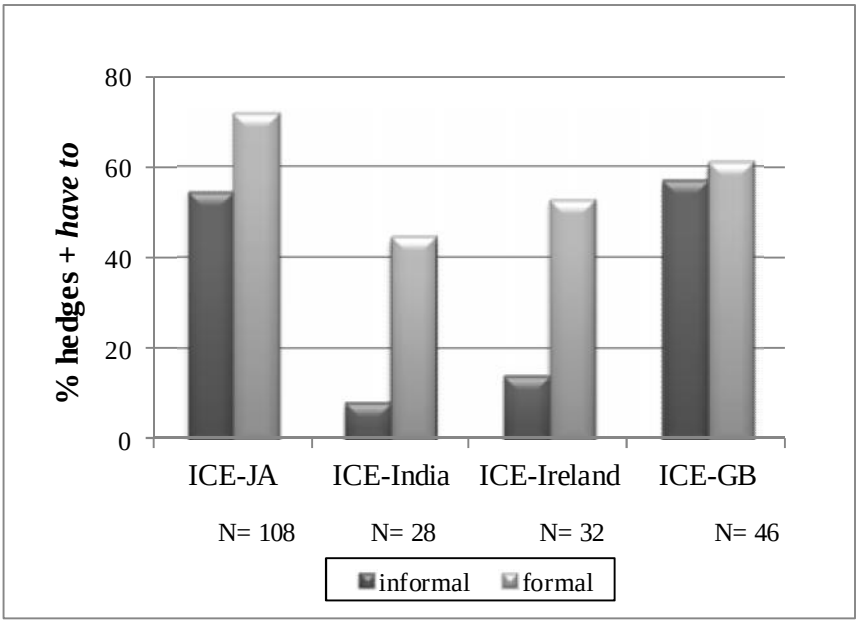

Figure 6.7. Distribution in percent of subjective *have to* with hedges

The figure above displays the distribution of such forms in percent in both formal and informal contexts in each dataset. Their distribution is relatively balanced across all the formal text categories in particular in ICE-JA, whereas a small number is found also in 'parliamentary debates' in ICE-India and in 'broadcast interviews' and 'broadcast discussions' in ICE-Ireland. It would appear then that such patterns are significant when accounting for modal use within outer circle varieties as compared to those belonging to

the inner circle. Thus, we may contend that the subjective element with this semi-modal is pragmatically motivated particularly in the dialogues from ICE-JA which is in stark contrast with the previous accounts which separate the meanings of *must* and *have to* in present-day English (see sections 4.5.3 and 6.3.3.2).

Next, it is useful to consider possible differences in the distribution of sources with *have to* as compared to *must*, which might implicitly offer more information on a possible specialization of meanings within New Englishes. Thus, a number of objective meanings occur in the text category of 'class lessons' in ICE-JA with 21 occurrences, ICE-India with 35 occurrences, and ICE-Ireland with nineteen occurrences. By contrast, I found only seven occurrences of objective *have to* in ICE-GB. These uses differ fundamentally from those with the modal, as the semi-modal here reports on general requirements which are related to an instruction or to specific circumstances. Hence, the semi-modal is less intense in meaning. Most of the times, *have to* correlates with generic second person *you* but also with *we*, denoting a discourse strategy to emphasise that the actualisation of certain actions hinges on external factors. Likewise, in some cases the semi-modal co-occurs with third person subjects referring to necessary processes related to an agent with which the speaker does not identify; (for comparison recall a similar example with *must* in (77)). In addition, such uses may co-occur with an inanimate subject as illustrated in example (98). A further text category which needs more examining is that of 'business transactions'. It is only ICE-JA and ICE-Ireland where objective uses of the semi-modal occur in large numbers. This observation appears to lend support to an earlier conjecture on the general distributional patterns with this semi-modal in the JamE dataset (see section 5.2.1). Moreover, unlike modal *must*, the semi-modal in these texts reports on necessary procedures in a negotiation between several participants.

Although semi-modal *have to* expresses more often discourse-external necessity, of more interest for the present purpose are the occurrences associated with discourse-internal information. Thus, roughly 59% of subjective *have to* in ICE-JA and ICE-India are found more often in the private dialogues unlike ICE-Ireland and ICE-GB where 50% and 57%, respectively, of such uses are preferred in public speech. The fact that both types of sources are equally distributed in the dialogues from ICE-Ireland is rather unexpected, however, offer more insights into the functional layering of this item in this

variety. The main observation is that subjective *have to* competes with subjective *must* at least in two text categories: such as 'broadcast interviews' and 'broadcast discussions. Further contexts contrast the occurrence of these items with a subjective element in ICE-Ireland, namely *have to* is found more often in 'class lessons', 'cross-examinations' and 'business transactions', whereas *must* is more preferred in 'parliamentary debates'. Overall, the same pattern for subjective *have to* is found in ICE-GB.

As with *must* (see section 6.3.2.4), subjective *have to* is compatible with speaker imposition of the obligation. Knowing that modal *must* was found to correlate with a discourse-internal source in the text category 'class lessons' I checked for similar uses with *have to*. It appears that only in ICE-JA does the obligation meaning with the semi-modal co-occur with definite second person *you*. This is borne out by the fact that the competition between the two items *must* and *have to* would be indicative for both colloquialisation and democratization of discourse. On the other hand, it might be also be due to the fact that the interpretation of subjective root reading with *have to* refers to "general or repeated requirements" (Westney 1995: 112), which acquires a more explicit meaning in the present tense. In this sense, the compelling situation in the interpretation of necessity with the semi-modal is also enhanced with the temporal information. However, this is then tantamount to the meaning of modal *must*.

To sum up, the examination of sources with *must* and *have to* offers, thus far, a contrastive picture of their use in spoken interaction in ICE-JA, ICE-India and ICE-Ireland. While with regard to modal *must* these varieties show more similarities with the British patterns found in ICE-GB, the data representing the two outer circle varieties also provide evidence for differences, and appear to be rather conservative. Intuitively, this might account for the speaker's different perception of necessities related most of the time to natural laws, rules or moral principles. Furthermore, at least in ICE-JA the use of *have to* might be linked to a tendency of dissolving inequality of powers in language. At the same time, one should not disregard a possible increased American influence of such modal behaviour. In the absence of comparable data from AmE it is difficult to make any strong judgements about possible changes in discourse in these varieties. However, while *must* is still quite well spread across the present data, at least in ICE-JA and ICE-India it tends to be associated with specific contexts of objective reading, whereas the semi-modal is on the way to becoming the default marker of

necessity. This pattern, then, would correspond to the general pattern found in present-day English.

Due to the dearth of data for *(have) got to* in ICE-JA and ICE-India, the discussion will resume with the main trends in ICE-Ireland as compared to ICE-GB. Even so, it is striking that only in ICE-Ireland the subjective-objective contrast with *(have) got to* is biased towards use in formal contexts whereas in ICE-GB this form is more preferred in informal conversation. In the previous description of sources with this item (see sections 6.3.1.3 and 6.3.2.3), I mentioned that *(have) got to* and its variant forms add emphasis, sometimes express urgency and even irresistible necessity, which is similar to the meaning of *must*. Most of the objective uses in the private dialogues in ICE-Ireland refer to a requirement or a specific circumstance, but may also express an internal disposition to the agent as in example (106), which I interpreted as being similar to the necessity with *need to*. In regards to public dialogues, most of the objective uses occur in 'class lessons', 'broadcast interviews' and 'broadcast discussions'. While in the first text category the speaker uses this item to express general requirements, in the latter two the necessity with this semi-modal is associated with the expression of urgency (see example (108)). The pattern for objective meaning with this item is similar to the one in ICE-GB. The majority of the forms co-occur with first and second person subjects particularly in informal dialogues and 'class lessons', but may occur also with third person subjects in both datasets. Interestingly, there is one instance in private dialogues from ICE-Ireland where more urgency of the necessity is achieved when the subject is inanimate and when the external factor is linked to compelling principles.

Finally, the evaluation of semi-modal *need to* yields interesting information about the extension to subjective uses in the New Englishes. To begin with, despite the contrast in their discourse frequency altogether, distribution according to level of formality suggests slightly different patterns for inner and outer circle varieties. Thus, discourse-internal sources occur in formal contexts in all datasets, whereas discourse-external sources occur mainly in informal contexts. Furthermore, in the private dialogues from ICE-JA, ICE-Ireland and ICE-GB *need to* is used with the basic meaning of internal disposition/ wish/desire of the speaker or an agent to the realization of an action and is most of the times preceded by first person *I*. By contrast, in the

private dialogues from ICE-India the objective meanings co-occur with a third person or an inanimate subject and denote unmarked necessity with the beneficiary of the action being unspecified or vague, and which according to Leech et al. (2009: 110) can be equally replaced by *have to*. In addition, the public dialogues include many instances which refer to general requirements such as an instruction usually in correlation with generic second person *you* and which, thus, appear as a more polite choice. As expected, these are found in 'class lessons' at least in ICE-JA, ICE-Ireland and ICE-GB. A further interesting occurrence was mentioned earlier in example (117) where the correlation with first person *I* was explained as being part of the internal flow of the discourse. There were two such occurrences in ICE-JA, one in the category 'class lessons' and another one in 'business transactions'. Overall, despite the fluctuation in frequency the use of objective *need to* in the public dialogues from ICE-JA, ICE-Ireland and ICE-GB converges to a great extent.

As already mentioned, it is the intrinsic value of unmarked necessity as combined with the subjective element of *need to* which involves pragmatic inferencing (see section 6.3.2.4). However, as the data shows, subjective *need to* may also express self-compulsion with first person *I* and personal opinion. The former are more typical for informal contexts whereas the latter for formal contexts. Moreover, the fact that at least in ICE-JA, ICE-Ireland and ICE-GB the most common type of subject with discourse-internal *need to* is impersonal *we* matches the observations from Smith (2003: 261) and Leech et al. (2009: 111) that the speaker identifies with the responsibility of the necessity. At the same time, such an obligation is projected as being in the common interest of all the participants in the conversation. Interestingly, there are at least two text categories which display preference for such indirect directives in these three datasets, namely 'class lessons' and 'business transactions'. While the two text categories were described as representing rather opposite formalities (see Figure 4.4), it would appear that the speakers are also more aware of the less face-threatening function of this item in interaction. Occasionally, these were found also in 'broadcast interviews', 'broadcast discussions' and in 'parliamentary debates'.

To conclude the discussion, the functional analysis shows that the behaviour of the four items in standard JamE, standard IndE and standard IrE largely converge with those found in the data from BrE. In addition, the evaluation according to text

categories and level of formality reveals further relevant information on the different 'ecologies' within root necessity, to use the term from Leech et al. (2009: 114). It seems, then, that register and level of formality can usefully account for the competition and layering within strong obligation/necessity. To return to the question which was raised at the beginning of section 6.3.1.1, the present analysis suggests that more than colonial lag, root *must* in ICE-JA and ICE-India might involve 'regressive divergence' (Hundt 2009: 15) whereas *have to* reveals extension of meaning to cover also subject-internal root necessity. This finding appears to be more representative for ICE-JA. By contrast, root necessity in ICE-Ireland is more similar to the data from ICE-GB. All in all, the more neutral marker *need to* seems to be a strong candidate for subjective root necessity competing with *must* and *have to* particularly in public speech whereas *(have) got to* is more typical for the two inner circle varieties.

6.3.3.2 Subjective root necessity in New Englishes: a case of ongoing change?

So far main attention has been given to the sources within root modality and the principal patterns in their distribution, this section though stresses the role of subjectification in the use of strong obligation/necessity in New Englishes. This is more challenging as usually one achieves a higher level of descriptive accuracy if semantic shifts are assessed in diachronic data whereas in the present case this phenomenon is based on synchronic data in the apparent time (see also discussion in section 1.3). Although such an approach seems less adequate for descriptions of regional and stylistic variation patterns, the semantic and pragmatic information incorporated in the subjective meaning can be accounted for in the light of subjectification. If this were the case, such tendencies would reveal more details about the grammaticalization of modal obligation/necessity in present-day English, hence, in New Englishes.

It has been assumed at several points in the course of the present analysis that the quantitative distribution of semantic contrast can be related to ongoing development in the use of *must*, *have to*, *(have) got to* and *need to* in the New Englishes under study. Perhaps the most remarkable finding is that the uses of both *have to* and *need to* are interesting cases of subjective necessity and are similar to previous findings on the evolution of other modal markers in present-day English (e.g. Goossens 2000; Krug

2000; Nordlinger & Traugott 1997; Traugott 1989; Traugott & Dasher 2002).[104] It is particularly the combination 'hedge (e.g. *I think*, *I believe*, *I feel*) + subject + *have to*' which needs closer attention (see section 6.3.2.2). To this end, it seems useful to refer to the earlier stages in the development of obligation *have to*.[105]

As Krug (2000: 97-102) convincingly argues, the origin of modalized meanings with this semi-modal is attested in the increased frequency of constructions of the type *have to say/tell*. The present case is, however, different for two reasons: first of all, *have to* is already modalized in present-day English and has become the default marker of root necessity in other varieties as well; and second, Krug's assessment focuses on an idiomatic expression which instantiates early stages of the grammaticalization of deontic/root meanings. Likewise, while the subjective root readings with modal *must* are associated with speaker authority the case of *have to* differs, as the imposition of obligation is inferable through a context-induced interpretation. The combination is thus pragmatically motivated, which reveals some similarities with the phrase *have to* and verbs of saying as claimed in Krug (2000: 101). In fact, this is an essential factor for the conventionalization of implicature, which grounds the root reading of the phrase in the speaker's attitude or belief to comply with a necessity (Bybee et al. 1994: 196; Traugott 1989). This is immediately apparent by the fact that the combination of the semi-modal with verbs of knowledge (e.g. *think*, *believe*) or perception (e.g. *feel*) places the root reading in the speaker's own evaluation of the proposition. In addition, there might be also a correlation with increased scope as discussed in Nordlinger & Traugott (1997). Nevertheless, the meaning is still of root necessity but which hinges on the speaker's support and evaluation of the situation on the basis of expert knowledge, and is therefore a subjectified instance (Goossens 2000: 164).

A further aspect to consider in the case of subjective *have to* is that it helps account for the intermediate steps in the semantic shift from root necessity to epistemic modality, and is similar to the "partially sanctioned" uses with modal *must* (see

[104] For a recent assessment of subjectification in cross-linguistic perspective in the evolution of present-day English *suppose*, which displays similarities with the semantic change of modals, see Visconti (2004).

[105] Although such combinations are common also with semi-modal *(have) got to* and *need to*, I consider the case of *have to* as different. As mentioned earlier, *(have) got to* is very close in meaning to *must* and denotes urgency of necessity whereas *need to* has the basic meaning of inherent need. Being often related to a general requirement the semi-modal *have to* may occur also with an ambiguous reading which warrants a more detailed examination of the various realizations from the perspective of subjectivity, hence, extension of meaning.

Goossens 2000). The transitional development presented here would correspond with a succession of several unidirectional intermediate steps of overlapping meanings (Heine et al. 1991: 113; Krug 2000: 101). Intuition would show then that root meanings with semi-modal *have to* preceded by hedges are appropriate contexts for the extension to epistemic necessity. From this perspective there might be a link between the overall high discourse frequency of *have to* in ICE-JA and ICE-India, the spread of subjective root meanings and the fact that these two datasets provide evidence for slightly more epistemic readings than the other two native varieties. Note, however, that the frequency analysis of semantic contrast did not yield statistical significance (see Figure 6.2). Moreover, in view of the aim to identify paths of change in New Englishes as well as achieve a general orientation in regard to the established standard norms of English, further aspects might prove just as relevant. Thus, all things being equal, the pragmatic component involved in this transition would suggest that speakers, in particular of JamE but also IndE, IrE and BrE, seem to be aware of the less face-threatening connotations of *have to* in combination with hedges which they employ productively in spoken interaction.

Regardless of the usefulness of such an account, subjectification is not necessarily limited to the cognitive processes in semantic change. Further language-external aspects should be considered (Narrog 2010: 408) which go beyond the explanatory power of frequency patterns. In this sense, the fact that the three varieties under study – JamE, IndE and IrE – have emerged under different socio-economical conditions might provide us with an answer for the layering of root necessity (see Chapter 3). As already mentioned, the spread of *have to*, *(have) got to* and *need to* in present-day English could be accounted for altogether as a result of a series of combined discourse-pragmatic and socio-cultural processes such as colloquialization, Americanization and democratization (see section 4.4.2). In the context of ex-colonial varieties such as JamE and IndE it would seem relevant to consider semi-modal *have to* and *need to* as exponents of more neutral and polite necessity as:

> Obligations potentially put human relationships at risk, since non-compliance is associated with all kinds of sanctions. Talking directly about obligations may be face-threatening or even menacing, and consequently, if the speaker is not in a position of full authority, puts the speaker her- or himself at risk. This is presumably even much more the case in small-scale, closely-knit, group-oriented societies, where the maintenance of human relationships is vital to the

> functioning of the group, than in large-scale, individualized cultures. (Narrog 2010: 409)

Even though one might assume stylistic factors (i.e. informal and colloquial use) to be responsible for the low occurrence of *(have) got to* in the two outer circle varieties, from the present findings it can be conjectured that its use is linked to the meaning of urgent and inescapable necessity which resembles authoritative *must*. As the corpus data shows, this is more evident in argumentative discourse in formal interactions. To conclude, more than the quantitative analysis, the assessment of the sources behind the necessity as related to subjectification provides valuable details about a possible reorganization within the domain of strong obligation and necessity in JamE, IndE and IrE.

6.4 The use of epistemic necessity

In order to emphasise ongoing evolution of these markers through subjectification, I shall move on to discuss the use of epistemic necessity in spoken data from ICE. As with root readings, the four modal markers display regional differences in the realization of epistemic readings (see section 5.2.1 and 5.2.2). Diachronic studies have identified epistemic modality as a development from deotnic/root modality towards greater subjectivity and evaluation of the meaning of the proposition (see Hanson 1987; Shepherd 1982; Sweetser 1990; Traugott 1989). Moreover, such an evolution is not only restricted to present-day English but appears to be a cross-linguistic, recurring phenomenon (Bybee et al. 1994). A cursory examination of the spoken dialogues will help support the theoretical assumptions about subjectification in epistemic necessity in data from ICE-JA, ICE-India and ICE-Ireland. For obvious reasons related to the low frequency of epistemic necessity in the present data, the discussion with this type of reading is less extensive than the previous one which focussed on root meanings. Although there is also a subjective vs. objective cline within epistemic necessity (see Table 4.2), the present section will not focus on such a distinction. Nevertheless, some preliminary remarks will provide a broader picture on the use of epistemic necessity in the New Englishes.

Typically, epistemic necessity is expressed with modal *must* as this marker is also most subjectified in present-day English. As the frequency analysis of semantic contrast has shown (see section 6.2), it is also this marker which most often expresses epistemic meaning in the present dataset as illustrated in the examples from (180) to (182):

(180) There **must** be different types of Indian Englishes because we have different varieties of <,> languages spoken (ICE-India.S1A-028.txt)

(181) I think it's both good and bad I mean <#>Well there **must** be some kind of bad in there you know because that's just how life is but generally I think that coming to UWI has prepared me for the world (ICE-JA.S1A.016.txt)

(182) Och your house is just so gorgeous now <#> You **must** be just so happy here <#> It 's really lovely (ICE-Ireland.S1A-006.txt)

To briefly exemplify the subjective vs. objective distinction within epistemic modality, instance (180) is a case of objective or logical necessity within epistemic reading which is conveyed by the speaker's confidence in regard to the conclusion that *there must be different types of Indian English* whereas the latter two (181) and (182) express subjective inferential meaning as the speaker's confidence or belief about the necessity of a situation. As in the case of root necessity there is a stylistic factor involved in the distribution by text categories with epistemic *must* as well. Thus, such meanings tend to occur more often in the private dialogues across all four corpora ICE-JA, ICE-India, ICE-Ireland and ICE-GB. It is in the formal text categories where the distribution varies in each dataset. While in ICE-JA these occur particularly in 'broadcast discussions' and 'business transactions', ICE-India is more like ICE-Ireland and ICE-GB, as such uses cluster mainly in 'broadcast interviews', 'broadcast discussions', 'parliamentary debates', 'cross-examinations'. In addition, some occurrences were found in the 'class lessons' in all the four corpora.

Often, epistemic readings express inferences about past situations and are more common with the modal *must* as conveyed with the perfect aspect (Coates 1983: 44) as in the examples from (183) to (187):

(183) And Kumata street you know that uh <,> <{1> <[1> yeah yeah<,> the airfreight has taken over <,> <{2> <[2> ah ha <,> they **must** have refused to <,,> take the envelop meant for ELPRO International <,> <{3> <[3>okay <,> and the postal

people **must** have just <,,> dumped it in the <,>undelivered parcels (ICE-India.S1A-094.txt)

(184) No <#>Ah</[></{> <$B><#>They **must** have just left the factory today (ICE-JA.S1B-071.txt)

(185) A few words in Marathi </O> and then you <,> you **must** have looked at the revolver <,> (ICE-India.S1B-063.txt)

(186) Now we would not be in the healthy situation that we are today and my sympathy to Deputy Quinn extends to the fact that I 'm sure he **must** have been quite livid that every morning he 'd get up for the past two weeks and read parts of what he was going to do in the budget in the papers (ICE-Ireland.S1B-060.txt)

(187) I just <,> I pulled out of the junction and I just went straight into the right-hand lane and I **must** have been in the right-hand lane a few minutes <,> five minutes (ICE-Ireland.S1B-061.txt)

All these examples are imbued with the speaker's certainty about a past action which is evaluated in the present. Thus, it is the non-realization of the action, such as the possible refusal to send an envelope, which is evaluated in the context of a telephone conversation in (183). By contrast, in (184) to (187) the speakers report on the speaker's confidence about the way an action has happened. A further particularity of epistemic readings is their co-occurrence with hedges as in (181), which was also illustrated in root readings. Here, hedges increase the subjective evaluation of the proposition obtaining pragmatic weakening.

As for the other three markers, the present data also provides evidence of their epistemic reading, however, exhibiting a higher degree of ambiguity. Both ICE-JA and ICE-India display epistemic *have to* as illustrated in the set of examples from (188) to (191):

(188) I guess in order to make sure that it functions properly they **have to** cut back on something to pay off their bills and keep themself <{5><[5>out of</[5> debt (ICE-JA.S1A-016.txt)

(189) And you **have to** go <,> to your branch very early I <{> <[> suppose </[> (ICE-India.S1A-092.txt)

(190) But I think India also won't let it go because <,,> uhm if India <w> let's </w> it go means it **has to** <,> let go also Punjab also then <,> (ICE-India.S1A-054.txt)

(191) So you had to turn your head to the side to see the shop <$A><#>From you come <indig>pan</indig> the corner you **have to** see the shop the corner is very deep <#>The car **have to** come on you to see if you are there <#>You **have to** come up to see the shop (ICE-JA.S1B-066.txt)

While the first three instances involve a subjective element with hedges, i.e. *I guess* in (188), *I suppose* in (189), *I think* in (190), denoting unambiguous epistemic reading (Krug 2000: 94-95) the last one expresses logical inference in reconstructing the events of an incident, such is typical for legal cross-examinations. Some further features favour an epistemic reading such as the conditional clause in (188) and inanimate *it* as well as the phrase *it means* in (190), which denotes inferred certainty about the predetermined course of an action. In this sense, both (188) and (190) would appear to be compatible with future reference (i.e. *it is necessary the case that they will have to cut back something / that India will have to let go*) of epistemic meaning which is inferable through an implicature. Notably, such realization would correspond to Bybee's et al. (1994: 240) path for necessity as illustrated in van der Auwera & Plungian (1998: 96) which suggests that epistemic meanings derive from future reference. However, it is questionable whether epistemic *have to* expresses temporal reference as such readings generally denote non-temporal inferences.

Instance (191) is noteworthy as the semi-modal reports logical conclusion about a past action, thus, is contra-factual. In addition, the semi-modal is uninflected for past reference and illustrates most likely an instance of a non-adapted local feature to the standard norm. This becomes obvious as the sequence also contains several indigenous words, most likely from Patois. It remains however unclear, whether epistemic *have to* in educated JamE is a creole influence and, thus, it may be an isolated instance. To sum up, the examples above suggest that there might be a link between subjective root and epistemic necessity with *have to* in order to explain the behaviour of this semi-modal as a case of subjectification particularly in the two outer circle varieties (see section 6.3.3.2). A further indicator for such a development is that epistemic *have to* also occurs in the formal text categories in all four datasets, however with the lowest number of occurrences in ICE-GB.

Finally, in spite of their low frequency, there is evidence for epistemic uses also with *(have) got to* and *need to*:

(192) Halle Berry'**s got to** be a really good actor (ICE-JA.S1A-015.txt)

(193) My plate is as full as it is now <#>Uhm but then if I see a situation where and you look around and there's no one else it's **gotta** be you<{10><[10><,></[10> and at that point I would be more than willing<{11><[11><,></[11> to step up to the plate so to speak <{12><[12>and uhm</[12> do my thin (ICE-JA.S1A-042.txt)

(194) It **has got to** be with the individual <,> and not with any government <,,> Because whether you are able to work or not <,> is something you can decide <,> (ICE-India.S1B-035.txt)

(195) Uhm apparently they **'ve got to** sign a documentation that <,> the sort of the ultimate veto of their copy <,> (ICE-Ireland.S1B-037.txt)

(196) But I guess<,></[2></{2> we **need to** have <}Although ><->a</-> <=>a</=></}><{1><[1><,></[1> degree debate one of these days <{2><[2>you know (ICE-JA.S1A-028.txt)

(197) Somebody **needs to** make the first move and normally I just tend to make the first move that's it (ICE-JA.S1A-051.txt)

(198) <S1A-060$B> <#> Stop moaning and get on with it
<S1A-060$A> <#> Well <{> <[> yeah </[> <#> But everybody **needs to** have a little moan on and off <,,> <&> laughter </&> <#> (ICE-Ireland.S1A-060.txt)

(199) If a man also wants to <,,> liberate himself then I think every man **needs to** liberate himself or <,> all human beings **need to** liberate themselves because <,> all of us have this basic impulse <,> to move towards some kind of an autonomy <,> growth <,> <{> <[> may be </[> this is something that we all <,> think about (ICE-India.S1A-079.txt)

Although epistemic *(have) got to* is associated with subjective necessity, it is only in (192) that such an interpretation becomes apparent. By contrast, (193), (194) and (195) denote logical conclusion that a situation is necessary. Except for (193) and (194) where the subject is inanimate, all the other instances correlate with animate subjects. As with the case of root meanings, epistemic readings with this semi-modal are more intense than, for example, with modal *must*. However, due to the low discourse frequency in the spoken dialogues from ICE-JA and ICE-India no clear tendencies in usage emerge. Nor can these be explained for ICE-Ireland as the epistemic readings in this dataset display the same level of frequency as the two outer circle varieties.

Further evidence for increased subjectivity in the data can be found with the few examples of epistemic *need to* in data from New Englishes, but which are less

transparent for an interpretation of inference. If (197) and (198) involve less explicit speaker-based inferences, the epistemic reading of (196) and (199) is identifiable because of the hedges *I guess* and *I think* which express the speaker's confidence. Despite the subjective element involved in such modalities, the extension to such meanings is still a minor feature of this semi-modal, which is apparent in the spoken dialogues in all the four datasets. Even so, the interpretation of the few epistemic examples hinges on conventionalization of implicature and thus converge with the findings in Nokkonen (2006: 63).

6.5 Summary

The present chapter has focussed on the uses and meanings of *must, have to, (have) got to* and *need to* in ICE-JA, ICE-India and ICE-Ireland. This analysis clearly shows that root modality dominates the uses of the four markers overall, both in the spoken as well as in the written press texts in these three varieties. The aim has been to identify the relationship between discourse frequency and the possibility of restriction or extension of meaning with these markers as well as to test on the basis of spoken dialogues the hypothesis, according to which subjectivity is associated particularly with *must* whereas objectivity with *have to, (have) got to* or *need to*.

The above argumentation departed from the assumption that the development from root to epistemic modality in English is characterized by several transitional uses which display increased subjectivity in utterance. The spoken data from JamE IndE and IrE show varying levels of subjectivity as conveyed with these markers. In addition it appears that there is a tendency of a gradual movement towards more subjective uses of *have to* and *need to* which differentiates the outer from the inner circle varieties and correlates with increased formality of the text categories. Tentatively, the asymmetry in the distribution of these markers is a consequence of the type of register and/or text category, respectively. An interesting finding refers also to the sharp contrast in the distribution of the root vs. epistemic distinction of the spoken dialogues and press texts from ICE-JA and ICE-India, which suggests that written language is more conservative.

Moreover, it appears that root *have to* in ICE-JA has extended to cover subjective necessity whereas in ICE-India and ICE-Ireland it is normally the case of modal *must*. Even so, it appears that the modal does not report prototypical subjective meanings but is used more often with the meaning of self-imposed obligation. Although the findings do not capture a clear pattern of either BrE or AmE influence on the three varieties under study, it might be that extension of meanings accompanied by the colloquialization of formal discourse are relevant factors in the spread of *have to* and *need to* in ICE-JA. By contrast, the findings from ICE-Ireland seem to converge with the pattern found in ICE-GB.

In regards to the subjective use with semi-modals, their interpretation often relies on pragmatic inference in combination with hedges which might be indicative of the subjectification of these items, and in particular of *have to*. In addition, it seems that apart from the semantic and pragmatic context, language-internal factors (e.g. subject type) account for the contrasting distribution of obligation and necessity even though this aspect was presented rather impressionistically. A systematic enquiry of the modal context on the basis of several language-internal and language-external factors will be the subject of Chapter 7, which is confined to a "descriptive-interpretative" approach involving statistical multivariate analysis.

7 Root *must* and *have to* in New Englishes: a multivariate analysis[106]

7.1 Introduction

The concept of variation in as a broad definition – i.e. covering synchronic (regional, social, stylistic) and diachronic variation – has been present from the outset in the present study. However, it is only in this chapter that a quantitative variationist approach, which has become standard in sociolinguistics investigations, will be employed. Although the present study is not a true sociolinguistic study (see section 4.3.1), it makes use of the variationist technique to investigate patterns of form-function variability which might be indicative of ongoing change (see Poplack & Tagliamonte 1996; 1999).

Even though the significance of functional distribution in the previous chapter (Chapter 6) could not be proven beyond doubt, a more fine-grained qualitative analysis has pointed towards several layers of more subtle facets of subjective and objective uses within root necessity in the three focus varieties. The outcome of semantic reading is, thus, a projection of the correlation between internal (grammatical) and external (explanatory) variables (e.g. regional stratification, text type).

The present chapter is restricted to assessing the variability of only the root meanings of the modal *must* and semi-modal *have to* because it is only within this type of modal reading that a set of distinct conditioning factors is identifiable for both markers. In addition, the total number of root readings with these two markers is large enough to permit the statistical analysis for this set of factors. Of particular interest here are the factors which condition the use of *have to* in the three New Englishes under study. Similar to sociolinguistic analyses, for the comparative assessment of regional and stylistic variation in the use of the two markers a functional equivalence was assumed between root *must* and *have to*.

In the following, I shall briefly outline the principles of the quantitative variationist approach in linguistic investigation, stressing its usefulness in the study of

[106] A previous version of this chapter can be found in Diaconu (2012). The present chapter is an improved version of this paper both as regards the assessment of root vs. epistemic contrast and the coding of the various explanatory variables with the logistic regression yielding slightly enhanced variance explained of the statistical model.

modal variation and change, and present the method underlying logistic regression in the present study (section 7.2). Next, I present the internal and external constraints and their distributional patterning across the data (section 7.3). Finally, I will discuss the statistically significant estimates in the light of ongoing change in the use of modal necessity in New Englishes (section 7.4). The chapter ends with a brief summary as well as some concluding remarks (section 7.5).

7.2 The variationist approach to language change

7.2.1 The importance of the linguistic variable

It has become relatively common in linguistic investigations to assess variation by means of statistical testing. Such quantitative approaches have emerged from an interest in the study of synchronic variation in non-standard varieties (for a review, see Bayley 2002), and are borne out by the observation that language change is reflected in systematic rearrangements of structural constraints as correlated with their frequency of use (Labov 1982: 75). In this sense, synchronic variation is often indicative of diachronic change.

Essentially, the variationist approach, also known as "descriptive-interpretative", investigates grammatical structures by means of multivariate quantitative techniques taking into account the linguistic form-function instability in discourse (Sankoff 1988: 141, 143). Relevant to the discussion is that such instability or variation is a characteristic of the functional equivalence between variant forms, and a primary condition underlying the evaluation of such variants is the heterogeneity of the speech community (Labov 1982). In contrast to univariate analyses the variationist approach builds on an empirical basis by/through the simultaneous assessment of linguistic constraints related to one form or the other, and which are evaluated against statistical significance and the relative strength of factors. However, for a better understanding of the speakers' linguistic choices one major task for the researcher is to first identify the constraints which will need to be considered/tested in a multivariate analysis. These constraints are usually a combination of language-internal and language-external factors (Tagliamonte 2002: 730).

7.2.2 Root modality and binary logistic regression

While quantification of variant forms of root necessity and their constraints has been undertaken on non-standard BrE dialects (Tagliamonte 2004; Tagliamonte & Smith 2006), on diachronic changes in both BrE and AmE (Jankowski 2004), as well as on other varieties, e.g. CanE (Tagliamonte & D'Arcy 2007), the present contribution extends the focus to include data from New Englishes.

The objective of such an analysis is to describe and quantify the effects of the relationship between different explanatory variables (e.g. grammatical, semantic, pragmatic and other language-external factors) which contribute to the outcome of a binary dependent/response variable (e.g. *must* of *have to*) in spoken data from ICE-JA, ICE-India, ICE-Ireland and ICE-GB. Such an approach will help complement the findings from the previous descriptive chapter (Chapter 6) as well as identify whether the conditioning effects are regionally specific "from which we can infer the structure (and possible interaction) of different grammars" (Tagliamonte 2002: 731). Extending this line of reasoning further, the regionally differentiated data might just reflect subtle grammatical readjustments (Tagliamonte & Smith 2006: 369) of obligation and necessity in JamE and IndE, which are expected to exhibit a different profile than IrE, and BrE.

Building on the assumption that gradience is an inherent feature of root necessity, logistic regression models are most suitable for the simultaneous assessment of factors contributing to the binary outcome (cf. Bod 2010: 635). Nevertheless, it should be borne in mind that the present assessment will not offer conclusive explanations for the observed differences in the frequency rates of these two markers, but rather will report the odds for the linguistic choice on the basis of probabilistic tools known from statistical analysis as:

> Logistic regression models estimate which of two outcomes is more likely to occur given that one or more independent variables (which may be scalar, categorical, or both) influence the outcome. (Szmrecsanyi 2006: 53)

Unlike many of the previous multivariate analyses which have modelled linguistic variation in the Varbrul suite of programs (see Bayley 2002; Sankoff & Labov 1979), the logistic regression in this study is carried out with R, an open-source programme for

sophisticated statistical measuring.[107] As the range of facilities is larger with such a general purpose package for statistical analysis (Johnson 2008: 74-75), the output will be also richer in information, such as the variance explained by the whole model including the interactions with internal predictors.[108]

The advantage of using logistic regression models consists in estimating coefficients for the degree of correlation between two or more explanatory/independent variables when trying to correctly predict the dichotomous outcome of a response/ dependent variable (see Geisler 2008; Gries 2008; Hinrichs & Szmrecsanyi 2007; Szmrecsanyi 2006). In this case the response variable is necessarily categorical as it incorporates the binary coding: "0" for *must*, and "1" for *have to*. Another fact to take into account regarding the response variable is that logistic regression produces estimates for the so-called baseline category coded as "0", which in this case is considered to be *must*. Similarly, the coding of each of the explanatory variables comprises this information. Each entry of *must* and *have to* from the spoken sections of the individual subcorpora from ICE displaying root readings (total N= 1823)[109] was manually coded for several explanatory variables. These are listed in Table 7.1 (see also discussion in section 7.3):[110]

[107] To this end, version R 2.8.1 for Windows was used. For more information, see R Development Core Team 2008 on the link www.cran-r-project.org.

[108] Particularly problematic in the case of the more common variable rules analyses is the lack of information on the overall variance explained by the model (Szmrecsanyi 2006: 53, 222 notes 229 and 211). On the limitations of multivariate analysis on modals and beyond using Varbrul, see also Bayley (2002: 130-132) and Kirk (1994).

[109] The root readings of *must* and *have to* incorporated in the present analysis are based on the occurrences found in ICE-JA, ICE-India, ICE-Ireland and ICE-GB which were analysed also in section 6.3.

[110] Bias according to type of medium (e.g. written vs. spoken discourse) was not included as an explanatory variable in this study, as the data comes exclusively from the spoken sections of each of the subcorpora of ICE.

Response variable:	*must* ("0") vs. *have to* ("1")
Explanatory variables:	
Language-external:	VARIETY ("BrE" vs. "JamE" vs. "IndE" vs. "IrE") TEXT TYPE (private dialogues "A" vs. public dialogues "B")
Language-internal: a) morphosyntactic:	GRAMMATICAL PERSON ("1pers." vs. "2pers." vs. "3pers.")
b) semantic:	REFERENCE ("generic" vs. "non_generic") VERB TYPE ("accomplishments/achievements" vs. "activity/states");
c) pragmatic:	SOURCE ("discourse_external" vs. "discourse_internal")

Table 7.1. Explanatory variables and their coding for the multivariate analysis

While the binary coding for the response variable is unproblematic, the baseline category refers to the predictors marked as "0": VARIETY=BrE, GRAMMATICAL PERSON=1 pers., REFERENCE=generic, VERB TYPE=accomplishments/achievements, and SOURCE=discourse-external (Geisler 2008: 39).[111]

Subsequent stages in the analysis consist of, first, creating a model by using *lrm* function[112] simultaneously for each variable in turn, followed by a stepwise regression in which we eliminate those variables that do not conform the significance level $p<.05$. In this way the elimination process contributes to testing and improving the model which is most adequate for our data (Geisler 2008: 38) as reported, for example, by the *Model L.R.* (log-likelihood-chi-square) estimation and the *Nagelkerke* R^2-value (Baayen 2008: 204; Hinrichs & Szmrecsanyi 2007: 459). While the first measure tests the statistical significance of the whole model, namely if the individual predictors contribute to explaining variance, the latter coefficient indicates accuracy in the proportion of variance ranging between 0 and 1 in the dependent variable and accounts for substantial significance if $R^2 \geq .05$ (Szmrecsanyi 2006: 54, 58). Next, the *odds*

[111] Within the two factor groups VARIETY and GRAMMATICAL PERSON several levels function as so-called *dummy* variables, also design variables (Geisler 2008: 39).
[112] The function *lrm* in R is the abbreviation for Logistic Regression Model.

ratio,[113] as well as *main* and *interaction effects* (Szmrecsanyi 2006: 58) of selected predictors will be taken into consideration. By measuring the effect of each explanatory variable it will allow for the indentification of similarities or differences with the modality systems from other (non-) standard varieties of English. From these one might extrapolate inferences about ongoing trends in obligation/necessity. Finally, I will report on the *predictive capacity* (%) of the model of accurately classifying the data (Gries 2008: 292). Ultimately, the data quantification provided by this analysis will allow us to infer predictions about the linguistic choice in spontaneous interaction, which are either diverging or converging in regionally and stylistically stratified data.

7.3 Operationalizing the variables

The explanatory variables in the present investigation (see Table 7.1) rely on criteria from the current literature which have proven to be useful in accounting for variability *within* root necessity. For the operationalization of these factors and assessment of the degree of embedding across JamE, IndE and IrE I shall combine Coates' (1983) cline of modality with Depraetere & Verhulst's (2008) classification (see also Chapter 6), which are able to jointly diagnose the synchronic behaviour of root necessity. As it turns out, these are comparable with factors identified in studies conducted by Tagliamonte (2004), Tagliamonte & Smith (2006), and Tagliamonte & D'Arcy (2007).

Recall, however, that this approach is slightly different from the previous studies enumerated above. Although the variable "speaker age" is not included in this assessment as would be the case in a true study in apparent time, regional setting and text type (informal settings of private conversations vs. formal settings of public conversations) will function as a sociolinguistic approximation, which will help situate the binary outcome at a particular stage of development. From this perspective, the study does not assess directly processes of change in time but the way each variety patterns with certain conditioning factors at a specific moment. This in turn will point towards reorganisation in the system, as a shift away from *must* towards *have to* (cf.

[113] In contrast to variable rules analyses which work with probabilistic weights, odds in logistic regression represent the ratio of the likelihood of an outcome (e.g. *have to*) occurring over the likelihood of it not occurring (Geisler 2008: 37; Szmrecsanyi 2006: 54).

genitive-choice in Hinrichs & Szmrecsanyi 2007: 439-440). Thus, the joint assessment of these variables helps to confirm or reject the hypothesis that they have a predicting effect on the use of the response variable (Jankowski 2004: 94).

While the discourse frequency and meanings of root *must* and *have to* has been discussed earlier in Chapter 6, the present section focuses on internal factors, such as the type of reference of the subject, grammatical person and the type of verb following the modal. It also looks at the way these combine with external factors such as, for example, variety or text type (i.e. private or public conversations).

7.3.1 Internal constraints defining the subject

According to Coates (1983: 36), an important parameter indicating the relative strength (strong vs. weak) of root modality is the category of subject, as encoded in the grammatical person.[114] For example, gradience on the cline from subjective to objective necessity is especially evident in second person, animate subjects (i.e. which denote speaker involvement in the utterance and exert authority over the subject; Coates 1983: 36). As it turns out from the total number of entries (N=1823) with root meaning, 86.2% (N=1573) of the subjects are animate, from which 68.1% (N=1242) occur with *have to.* Interestingly, inanimacy is almost equally distributed between the two expressions even though a visibly smaller rate: 6.3% (N=106) of inanimate subjects occur with *must*, whereas 7.8% (N=144) favour *have to*. Given that root modality most often encodes speaker imposed necessity which compels an agent, usually human, to undertake an action, the fact that the majority of subjects preceding *have to* are animate suggests that this marker is deeply embedded within the wider domain of root necessity (Krug 2000: 90).

In regard to the variable GRAMMATICAL PERSON the two expressions vary consistently in our data. As Tagliamonte (2004: 46) shows, there is a close link between grammatical person of the subjects and type of reference, that is whether these report on a generic or definite subject. Figure 7.1 collapses the data together on subject type and

[114] Analogously, in the functional linguistics framework the subject preceding the modal encodes *speaker*- and *content-relatedness* of modality, and which is also closely linked to the distinction between subjective and objective modality (Verstraete 2001: 1506).

reference in percent from four different corpora: ICE-JA, ICE-India, ICE-Ireland and ICE-GB. Interestingly, *have to* clusters especially around first person definite with 35.4% (total N=962 of definite subjects) and second person generic subjects with 38.6% (total N=861 of generic subjects). *Must* is preferred for first person with 9% and third person definite subjects with 12.6%, and less with second person subjects (regardless of the type of reference), which supposedly confers obligation meaning to the modal reading. From this distribution it emerges that *have to* might be taking over the contexts commonly assigned to *must*. By contrast, the modal appears to recede in contexts which were assumed to be typical for the semi-modal, and which is in contrast with data from non-standard dialects presented in Tagliamonte & Smith (2006: 361).

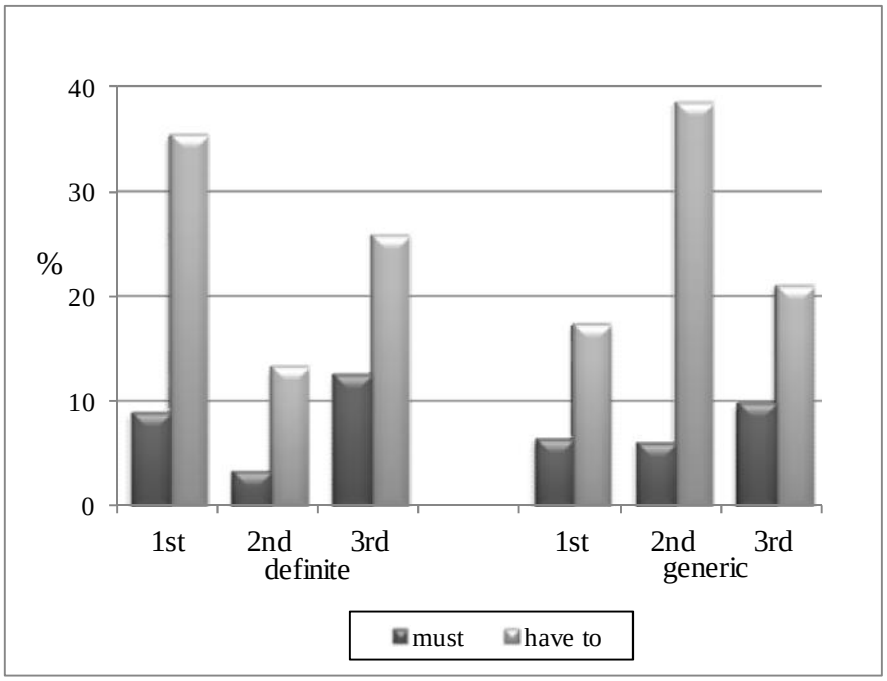

Figure 7.1. Distribution in percent of grammatical person according to reference of the subject in spoken dialogues from ICE-JA, ICE-India, ICE-Ireland, and ICE-GB

Even though the figures above are not presented as related to their distribution in the individual varieties we may conjecture that they incorporate the recent changes observed elsewhere in present-day English.

Note that the distinction made by the subject referent is directly reflected in the strength of necessity. That is, definite readings as in (200) and (201) are considered as

indicators for strong obligation/necessity, whereas in (202) and (203) the generic subjects indicate weaker necessity:

Definite readings:

(200) Yes of course I will come sometimes if god sends me <,> and you too **have to** come (ICE-India.S1A-065.txt)

(201) I wish I had time to uhm to do it myself but I **have to** be at work and all of that so (ICE-JA. S1A-008.txt)

Generic readings:

(202) We're obsessed with this idea of the weather as well and you kind of feel that you **have to** comment upon the weather with everybody you meet (ICE-Ireland.S1B-004.txt)

(203) And as I always say there 's no need for anybody to starve in this country and if your pride won't allow you ask <,> it 's a false pride <,> <#> You **must** be prepared to get up and ask if you need food (ICE-Ireland.S1B-040.txt)

By and large, Figure 7.1 shows similarities with the patterns found in Tagliamonte & Smith (2006: 360), however, regarding the preferences of *have to*, this data is in contrast to their results on the eight British dialects from Scotland, England, York, Northern Ireland and Sussex. This is because in addition to first and third person definite contexts which favour *have to*, there is a third layer in our data, namely that of second person generic subjects which correlate with the semi-modal. This, in turn, suggests that *have to* has extended its use to further contexts, i.e. those related to speaker imposed necessity.

7.3.2 Verb type

A further constraint which accounts for the difference between root and epistemic modality refers to the verb following the modal (Coates 1983: 33). Hence the differentiation according to lexical aspect could be a sound indicator for the different regional uses of necessity in ex-colonial varieties of English.

In the following, we consider lexical aspect (*Aktionsart*) of the verb preceding the modal as an indicator of strength on the cline from strong to weak necessity.

Following Vendler (1967: 97-121), a two-way distinction was adopted, which comprises verbs expressing accomplishments (e.g. *paint*, *build*) or achievements (e.g. *recognize*, *notice*) and states (e.g. *know*, *love*) or activities (e.g. *run*, *drive*), respectively. The motivation for this two-way distinction is based on telicity (Huddleston & Pullum 2002: 120), namely that verbs of accomplishment/achievement have an endpoint (i.e. they are telic) whereas activities/states do not have an endpoint (i.e. they are atelic). This classification differs from the one in Quirk et al. (1985: 201) which separates activity or dynamic (including 'durative' and 'punctual' verbs) and stative verb meanings. Therefore, any comparison with other studies which include the dynamic vs. stative distinction, e.g. Tagliamonte (2004),[115] is rather limited.

As a rough generalization, we may assume that strong root readings occur when the modal or semi-modal collocates with an accomplishment/achievement verb as in (204) and (205), whereas activities/states would indicate weak reading as in (206) and (207).

Accomplishments/achievements:

(204) In order to communicate you **must** learn my language because you have a non-language (ICE-JA.S1A-003.txt)

(205) Not just because <w> I'm </w> I A S <w> I'm </w> appointed as a teacher I **have to** go there <,> teach there <,,> and teach them for thirty thirty-five minutes whatever is in the book you know <,,> (ICE-India.S1A-003.txt)

Activities/states:

(206) So I don't have to plan to get married because <,> I **must** wait for the suitable person <,> who is for me (ICE-India.S1A-031.txt)

(207) Yeah <,> <w> it's </w> a residential school and I **have to** stay throughout the year (ICE-India.S1A-022.txt)

Figure 7.2 displays the overall distribution in percent of the verb type following either *must* or *have to* combining all the four datasets of interest.

[115] The classification adopted in Tagliamonte (2004: 44-45) is between stative/durative vs. punctual verbs. According to this distinction, strong root readings are favoured with punctual verbs whereas weak readings with stative/durative verbs. However, the criteria for such a distinction remain unspecified.

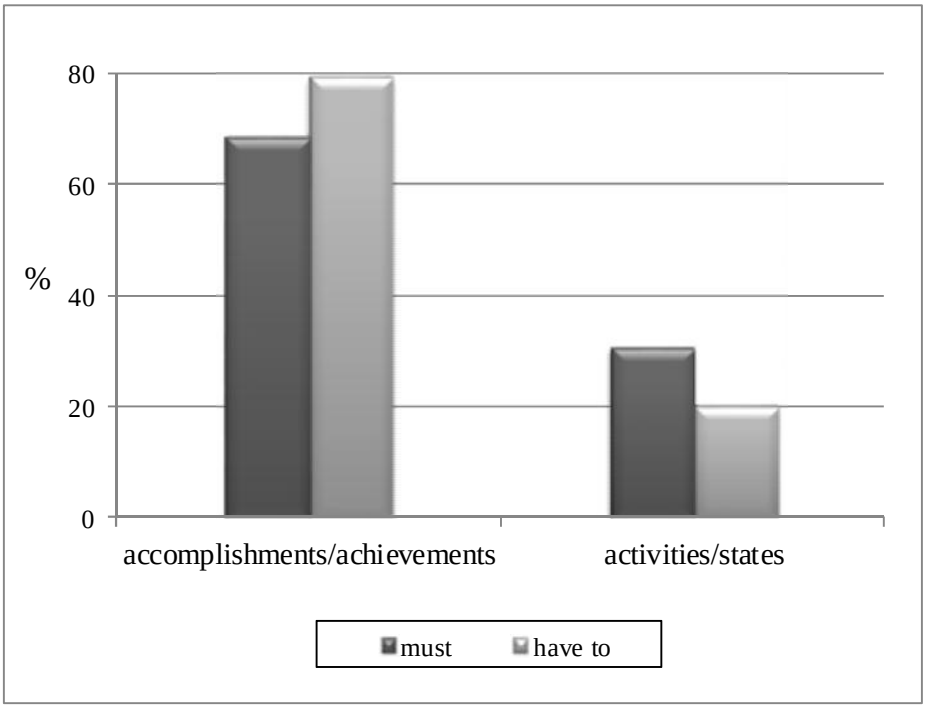

Figure 7.2. Distribution in percent of verbs expressing accomplishments/achievements and activities/states in the spoken dialogues from ICE-JA, ICE-India, ICE-Ireland and ICE-GB

For each variant form, the distribution according to accomplishments/achievements and activities/states, respectively, adds up to 100 percentage points. Again, the ICE database suggests stability in the overall distribution of verbs where the majority of verbs expressing accomplishments/achievements occur with *have to* (79.6% from a total N=1208 of verbs expressing accomplishments/achievements) as opposed to *must* (68.9%). In contrast, activities and states are most preferred with *must* (31% from a total N=615 of verbs expressing activity/states), and are closely followed by *have to* (20.3%). At this point it is useful to draw a parallel, an albeit limited one, to the findings from British dialects (Tagliamonte 2004: 44), where this semi-modal is more likely to be followed by stative verbs. It follows that the present data reveals a different configuration of internal constraints than suggested for other varieties. Although there is no evidence yet for statistical significance, the variable VERB TYPE could be a possible factor indicating that changes are under way in this particular area.

7.3.3 Internal factors: outer vs. inner circle

As mentioned earlier there are variable functional contexts for the use of modal *must* and semi-modal *have to*. Therefore, it will be instructive to extend their distribution in the four regional datasets under study and to identify underlying patterns of reorganisation by means of one-dimensional analysis.

Tables 7.2 and 7.3 display the distribution in percent of generic and definite subjects, respectively, with both *must* and *have to* across the spoken conversations from ICE-JA, ICE-India, ICE-Ireland and ICE-GB. The percentage is calculated from the total amount representing both subjects for *must* and *have to*. The two tables reveal interesting regional distributions according to the type of reference and grammatical person. Despite the different quantitative preferences in the outer and inner circle varieties, these are balanced within the four datasets.

%	**Generic subjects (N=861)**					
	must			***have to***		
	1.pers.	2.pers.	3.pers.	1.pers.	2.pers.	3.pers.
ICE-JA	2.3	2.4	3.2	17.4	4.6	2.9
ICE-India	3.6	2.5	5.1	7.5	11.4	6.3
ICE-Ireland	1	0.6	0.5	8	0.6	2.2
ICE-GB	0.3	0.3	6.9	5.6	0.6	2.3
Total	**7.1**	**5.8**	**15.7**	**38.5**	**17.2**	**13.7**

Table 7.2. Distribution in percent of generic subjects with *must* and *have to* according to GRAMMATICAL PERSON and VARIETY

%	**Definite subjects (N=962)**					
	must			***have to***		
	1.pers.	2.pers.	3.pers.	1.pers.	2.pers.	3.pers.
ICE-JA	1.8	1	1.9	11.6	2.8	6.6
ICE-India	2.1	1.1	5.7	10.6	4.9	9.2
ICE-Ireland	2.2	0.4	1.9	8.7	3.4	6.1
ICE-GB	2	0.8	2	4.4	2.1	3.9
Total	**8.1**	**4.7**	**11.5**	**35.3**	**13.2**	**25.8**

Table 7.3. Distribution in percent of definite subjects with *must* and *have to* according to GRAMMATICAL PERSON and VARIETY

In addition, such differences also occur within the two outer circle varieties, although to lesser degree, for example ICE-India exhibits a higher proportion of generic subjects

with *have to* as compared to the other datasets. ICE-JA displays an even higher proportion of first person generic subjects with *have to*.

Additionally, Table 7.3 shows that, despite minor differences, the four varieties under study largely converge with definite subjects, in particular with *must*. It is with the semi-modal *have to* where first person and third person definite subjects are more preferred in ICE-JA and ICE-India. Nevertheless, the overall picture is of stability and, therefore, it is difficult to account for shifts in usage at this level of analysis.

Table 7.4 below lists the distribution of both *must* and *have to* in percent according to lexical aspect of the verb in each subcorpus in ICE. The percentage is calculated from the total number of each accomplishments/achievements and activities/states verbs separately.

%	**Accomplishments/Achievements (N=1208)**		**Activities/states (N=615)**	
	must	***have to***	***must***	***have to***
ICE-JA	5.5	25	8.9	21.9
ICE-India	8.4	29	14.3	22.6
ICE-Ireland	3.3	16.1	3.7	15.2
ICE-GB	2.9	9.3	4	9.1
Total	**20.1**	**79.4**	**30.9**	**68.8**

Table 7.4. Distribution in percent of modals according to VERB TYPE in the spoken dialogues ICE-JA, ICE-India, ICE-Ireland, and ICE-GB

The regional distribution reveals that the highest percentage rates cluster in ICE-JA, ICE-India and ICE-Ireland around both dynamic and stative verbs following *have to*. Moreover, it is only ICE-India which shows the highest preference for stative verbs following modal *must*. Surprisingly, ICE-GB shows an overall lower rate of distribution according to verb type in particular with the semi-modal. In order to interpret the relative contribution of these factors as a shift in the constraints selected either by the modal or semi-modal, it is necessary to identify which of these provide underlying, statistically significant information on variability.

7.4 Multivariate analysis: results and discussion

7.4.1 Contribution of individual predictors

Previous studies on modality have applied logistic regression models by simultaneously measuring the effects of a number of factors on the outcome of the response variable. The same procedure was used in the present case. Thus, in a first step I performed a logistic regression with the variables found, for example, in Tagliamonte (2004), Tagliamonte & Smith (2006), and Tagliamonte & D'Arcy (2007), extending it subsequently to correlations/interaction terms with four 'focal' variables: VARIETY, TEXT TYPE, REFERENCE and SOURCE (see section 7.4.2). Using this particular technique means accounting for the variance of the binary response variable by obtaining the most adequate model at the significance level $p<.05$.

The following is a summary of the most adequate model accounting for individual contributions to predict the response (see Table 7.5). The first observation is that significance of the whole model is provided by the log-likelihood value, whereas the predictive efficiency of correctly classifying each construction is roughly 75%. The very small *p*-value at 8 degrees of freedom of the model log-likelihood chi-square value indicates that the predictors are overall significant (Baayen 2008: 204).

Explanatory variables	**P**	**Odds ratio**
Intercept	.0000	3.32
VARIETY=IrE	.0404	1.58
TEXT TYPE=B	.0000	0.45
VERB TYPE=activities/states	.0001	0.65
GRAM. PERS=2 pers.	.0366	1.38
GRAM. PERS=3 pers.	.0045	0.67
SOURCE=discourse-internal	.0001	0.61
Model L.R.=143.6; $p \approx 0$ at 8 d.f.		
Nagelkerke R^2=0.113		
predictive capacity 75%		

Table 7.5. Output of significant estimates predicting the odds for *have to*

Moreover, the output suggests that several individual predictors are involved in the choice for the response variable. However, the valued displayed by *Nagelkerke* R^2=0.113 is rather low accounting for only 11% of the variance explained by the model.

This means that other predictors might have been missed or that variance is due to free variation (Hinrichs & Szmrecsanyi 2007: 460). Although it can be considered substantially significant (≥5% variance explained) the low coefficient does not allow for confident assumptions about the variance accounted for in the independent variables. Under these circumstances the model will be evaluated as indicating tendencies for the behaviour of root necessity in JamE, IndE and IrE.

Note that the estimates and odds are given for the construction coded as "1" which in this case is the semi-modal *have to*. By contrast, Intercept in Table 7.2 marks the estimates and odds for the baseline categories coded as "0" (Geisler 2008: 39) and which jointly yield statistical significance at $p<.0000$.[116] In regards to the main effects of each of the variables, the negative coefficient with a factor 0.46 displayed by the predictor "TEXT TYPE=B" indicates instead a reverse (positive) effect as associated with modal *must*. In other words, formal contexts lower the probability of the choice of the semi-modal.[117] This is not a surprising outcome, as it was already made clear from the distribution of these expressions that *have to* occurs more in informal settings. At the same time, this finding is a further confirmation that *must* is the more formal choice in spoken interaction.

Similar results are obtained with the predictors "VERB TYPE=activities/states" and "SOURCE=discourse-internal" which have a negative effect on the odds for the semi-modal *have to* reducing the odds by factor 0.67 and 0.62, respectively. A positive effect of increased odds indicates that modal *must* is favoured. The fact that verbs expressing activities and states increase the odds of using *must* suggests that the root meaning of this modal is restricted to specific contexts. This kind of modal behaviour is different though from the evidence found in non-standard English, e.g. Tagliamonte (2004: 48), where this modal was found to be favoured with punctual verbs.

Regardless of the fact that discourse-internal sources lower the odds for the choice of semi-modal *have to*, the present output suggests that pragmatic factors have a

[116] Values which highlight main and interaction effects for ICE-GB can be obtained through conversion of the binary coding ("0" or "1") for the factor levels in the variable VARIETY. Thus the codification depends exclusively on the personal evaluation of the researcher.

[117] Visualization of odds as percentages is achieved by subtracting 1 from the odds and multiplying by 100 (Gries 2008: 289). For example, the percentage change in the odds for the variable "style=B" are calculated as (0.4657195 -1) ·100= -53.4%; where 0.4657195 represents the main effect of the explanatory variable which influences the outcome of the response variable. The main effects can be transformed as factors, in this case as factor 0.48.

predictive effect on the outcome of root necessity, in general. Conversely, significant estimates pointing to negative effects deserve reporting as they may indicate an incipient tendency in the use of semi-modal *have to* towards new grammatical environments, though displaying low statistical probability. Related to this issue, it will be useful to return to the earlier discussion on the subjective vs. objective opposition from the previous chapter, where it has been suggested that semi-modal *have to* is currently extending to express also subjective necessity in present-day English.

Next, it turns out that the regional factor IrE as well as the grammatical person, specifically, second person subjects, yield the strongest odds for predicting *have to* as the response.[118] This is a remarkable finding, as it was expected that an IndE or JamE speech setting would greatly increase the odds that the semi-modal is the response. More than that, the two varieties were not selected as statistically significant. Likewise, no statistically significant information is provided for variable REFERENCE. On the other hand, the regional effect of IrE confirms previous findings on this variety (Corrigan 2000; Tagliamonte 2004; Tagliamonte & Smith 2006).

So far, six predictors from five factor groups produced significant estimates though, mainly for the case of modal *must*. This is broadly consonant with what has been hypothesised in the literature about strong obligation and necessity. Except for the regional effect found for IrE and of second person subjects, the regression estimates that verbs expressing activities/states, third person subjects as well as discourse-internal sources raise the odds for the likelihood to use *must*. More striking is that such a configuration expresses less speaker imposed obligation, hence weak obligation (cf. Coates 1983: 37), which corresponds thus with the recent changes in the use of this marker (see Depraetere & Verhulst 2008). Note also that the same observation was made – though at a one-dimensional level of analysis – in the discussion on semantic contrast according to subjective and objective necessity (see section 6.3.3).

To sum up, it appears that the estimates for these predictors (semantic: VERB TYPE and SOURCE; grammatical: GRAMMATICAL PERSON; and language-external: TEXT TYPE) are most accurate when accounting for modal alternation. Furthermore, *have to* in our data correlates less with activities/states verbs than *must*, as the negative effect

[118] I refer to a strong effect if the coefficients above in Table 7.2 deviate from 0 and do not equal 1 (Gries 2008: 289; Szmrecsanyi 2006: 56).

shows. Likewise, formal speech settings and subjective sources are more likely to predict modal *must* instead of *have to*.

7.4.2 Interaction terms

Despite obtaining statistically 'non-significant' values for some of the factors it may be that these contribute to significant interactions with other explanatory variables, and increase the amount of variance explained of the model. Therefore, these non-significant values should not be eliminated from the model. Further indications possibly pointing to new trends within obligation and necessity expressed with *have to* in the three varieties – JamE, IndE, and IrE – can be disentangled by looking at shifts displayed by the odds ratio of the interaction effects. What the model should capture, then, is "where the effect of an independent variable (the 'focal' variable) on the outcome differs depending on the value of a third variable a so-called 'moderator' variable" (Szmrecsanyi 2006: 57). In other words, the quantification of the variables which are identified as contributing, for example, to the subjective vs. objective dichotomy of the response variable could be indicative for reorganization in the system, either in the varieties pertaining to the outer circle or inner circle. From this perspective, interactions can act as an improvement or as an adjustment of the predictors to the model.

The methodological steps consist of maintaining all the variables by setting contrasts between each of them in turn and generating a new object every time (Gries 2008: 287). Thus the logistic regression was processed identically, four times with the chosen 'focal' variables.

In regards to the qualitative improvement of the model, note that the overall variance explained by the model increases in comparison with the regression on the individual contribution of the factors with R^2-values ranging from 0.127 to 0.165. Although this is still not an optimal result, overall the model accounts for up to 16% of substantial significance in variance explained. Notwithstanding, the predictive capacity of correctly classifying the data is maintained at 75%, which is similar to the capacity assessed in the case of main effects.

As becomes apparent from Table 7.6, introducing interaction terms produces changes in the predictive effects of the 'focal' variables (Hinrichs & Szmrecsanyi 2007: 464), which will be interpreted in terms of the restructuring of root necessity within New Englishes. By adding interaction terms, clear strings of constraints can be seen, which in the regression on individual predictors could not be identified.

Focal variable	**Interaction terms**	**P**	**Odds ratio**
VARIETY	VARIETY=IndE * TEXT TYPE=B	.0003	0.27
	VARIETY=IndE *REFERENCE=non_generic	.0016	4.01
	VARIETY=JamE * REFERENCE=non_generic	.0129	3.04
	VARIETY=IrE * SOURCE=discourse-internal	.0194	0.36
TEXT TYPE	VARIETY=IndE * TEXT TYPE=B	.0003	0.26
	VARIETY=IrE * TEXT TYPE=B	.0421	0.41
	TEXT TYPE=B * SOURCE=discourse-internal	.0001	2.54
SOURCE	TEXT TYPE=B * SOURCE=discourse-internal	.0021	2.12
	GRAM. PERS.=3 pers. * SOURCE=discourse-internal	.0000	4.98
	GRAM. PERS.=2 pers. * SOURCE=discourse-internal	.0071	2.63
REFERENCE	VARIETY=JamE * REFERENCE=non_generic	.0079	3.10
	REFERENCE=non_generic * SOURCE=discourse-internal	.0028	0.47

Table 7.6. Selected interaction terms and output of significant estimates predicting the odds for *have to*

To begin with, the regional effects of interaction terms on the likelihood of *have to* reveal a contrastive pattern. The regional effect in the interaction VARIETY=IndE * TEXT TYPE=B produces a change of odds ratio in the main effect of the focal variable by the multiplicative factor 0.27 (Hinrichs & Szmrecsanyi 2007: 465). In other words, the

regional effect of IndE ($0.69 \times 0.27 = 0.18$)[119] lowers the odds by 0.18 for the likelihood of the semi-modal *have to* in formal speech contexts. This also means that the main effect of IndE on the likelihood that *have to* will be favoured in text type B is different than in the case of text type A. Recall that earlier significant regional information was offered only for factor level IrE. It appears from the interaction that this regional factor is not a particularly strong predictor of odds conditioning semi-modal *have to* in such contexts.

A similar effect is found in the interaction terms VARIETY=IrE * TEXT TYPE=B. However, in this case it is the main effect of text type B on the outcome (*have to*) which changes in an IrE context, lowering the odds by a factor 0.41 as compared to another regional coding comprised within this factor group. Unsurprisingly, formal speech settings have a less favouring effect on semi-modal *have to* overall. In addition, we find the same disfavouring stylistic effect with IndE in the interaction VARIETY=IndE * TEXT TYPE=B, this time by a factor 0.26., which can be related to a slower pace of the 'colloquialization' of this marker in confined speech settings.

Similarly, the interaction VARIETY=IrE * SOURCE=discourse-internal suggests that the effect of IrE on the semi-modal *have to* is less strong in contexts where the necessity has a discourse-internal source. Such an outcome confirms to some extent the previous assumptions about the behaviour of this semi-modal as correlated with the source of necessity. However, recall that this does not mean that *have to* cannot express subjectified meanings, only that within this particular regional configuration the odds are lower. A more appropriate interpretation might also be that subjectivity, hence, subjectification of root *have to*, is still in an incipient phase. From this perspective, the present assessment does not invalidate the qualitative evaluation made in section 6.3.3.1, but rather offers more fine-grained information on the predictive capacity of this constraint. Moreover, such effects might be indicative of a more general trend currently observable in present-day English, that the semi-modal *have to* is not necessarily a direct substitute for modal *must*, as it stands for a more neutral or less direct marker of speaker authority.

The last disfavouring effect was obtained in the interaction terms REFERENCE= non_generic * SOURCE=discourse-internal which means that the main effect of definite

[119] The main effect of the factor level VARIETY=IndE is 0.69.

subjects lower the odds by factor 0.47 for the semi-modal as the choice if the necessity denotes a discourse-internal source. Conversely, a change in the effect of definite subjects on *have to* as the linguistic choice is obtained when the necessity reports on objective/discourse-external sources. This is a surprising effect, as the literature on root modality (Coates 1983) states that definite subjects indicate speaker involvement, which is usually related to subjective necessity. Again, the interaction term confirms that reference of the subject is not an important predictor of root necessity in the dataset and is in stark contrast with other studies, e.g. on non-standard BrE dialects (Tagliamonte & Smith 2006: 366, 368) and on CanE (Tagliamonte & D'Arcy 2007: 78) which suggest that *have to* is favoured by definite reference and objective necessity.

Nevertheless, strong regional effects are recorded for IndE and JamE, respectively, in interaction with definite/non-generic subjects (see Table 7.6). Recall also the frequency analysis of the distribution according to reference of the subject, which revealed that apart from first and third person definite subjects, *have to* also collocates with second person generic subjects (see Figure 7.1). Likewise, the distribution according to definite subjects in ICE-JA and ICE-India (see Table 7.3) showed slight differences compared to ICE-Ireland and ICE-GB. In the light of these interactions the regional main effect increases the odds for the semi-modal, particularly, with definite subjects as compared to generic subjects (with factors of 4.01 and 3.04, respectively). A further corroborative strong effect is evident when definite subjects interact with the factor JamE (multiplicative factor 3.10), which indicates that this variable is an important predictor for modal choice, and in particular in this variety after all.

Finally, there are three different interaction terms which are variety-neutral, but which suggest that the pragmatic factor of SOURCE has a strong effect on the outcome in formal contexts, both in the second and third person. Although regional factors did not yield favouring effects in correlation with this pragmatic variable, it appears that this interaction is regionally invariant, which can be a sign of "ongoing re-weighting of internal grammatical constraints" (Tagliamonte 2004: 52).

Contrary to expectations, the interaction GRAMMATICAL PERSON=3 pers. * SOURCE=discourse-internal reveals that the effect of discourse-internal sources on the outcome changes by a factor of 4.98 when there is a third person subject involved as

opposed to the other categorical codings. Recall that earlier the individual contribution of the factor representing grammatical person produced a negative effect on the odds predicting *have to*, which was interpreted as favouring the modal *must*. Although the interaction term does not involve information on the reference of the subject it is comparable to a similar effect in non-standard dialects, where first person and third person definite subjects correlate with subjective necessity (Tagliamonte & Smith 2006: 368).[120] Such an outcome is consonant with the previous assumption (see section 6.3.3.1) that semi-modal *have to* is a more suitable option of expressing strong necessity in JamE.

Overall, the analysis shows that the linguistic choice in the expression of modal necessity in New Englishes appears to be conditioned by an underlying set of regional and pragmatic factors.

7.5 Summary and conclusion

The present chapter has applied the variationist method for the statistical modelling of various predictors assumed to constrain the use of *must* or *have to*. Although the model itself has proved insufficient to fully explain overall variance, the several significant values are indicative of some tendencies emphasised earlier in this study, as well as in other corpus-based studies.

The use of modal *must* and semi-modal *have to* as expressions of root obligation/ necessity has turned out to correlate with several conditioning factors. The several variables which I analysed, whether individually or jointly, offer a rather stable picture in regards to the alternation of the two expressions in dialogues from IndE, JamE and IrE. In several respects the detailed effect of these factors in the three New Englishes is different from other regional varieties: e.g. British dialects (Tagliamonte 2004; Tagliamonte & D'Arcy 2007; Tagliamonte & Smith 2006).

Although the univariate analysis of *must* and *have to* did not yield significant differences between the inner and outer circle, in the regression which measured the

[120] Note, however, that their approach is slightly different with regard to interaction terms, as their study is based on Varbrul.

individual contribution of factors two out of five predictors (i.e. regional factor IrE and second person subjects) yield positive effects on *have to* as the response. Text type (here 'formal'), verbs expressing activities/states and discourse-internal sources of obligation correlate positively with modal *must*. This suggests that root readings with *must* in our dataset are likely to occur with specialized functions, suggesting a very late stage of development.

A further aspect in the multivariate analysis reveals that the regional factor is by itself not necessarily a strongly predictive variable in the case of root *have to*. This is not entirely surprising as *have to* has been on the rise in all varieties of English since the eighteenth century. Adding interaction terms to our statistical model shows that speech community as a variable has a significant effect, most of the times disfavouring the semi-modal in correlation with specific stylistic and pragmatic constraints. Nevertheless, the regional factors of JamE and IndE have a favouring effect on the semi-modal when they interact with semantic variables.

Despite the significant stylistic, pragmatic or semantic effects in the data, no conclusion can be drawn that visible shifts are under way in the use of *have to*. Nor can a straightforward answer be found regarding the initial explanations for the differences found in the data. Evidence more or less supporting the hypotheses of 'colloquialisation', as well as 'democratisation' could be reflected by the interactions which favour the likelihood of *have to* in contexts involving formal texts, generic subjects and discourse-internal sources. In addition, regardless of regional context, the statistical evaluation suggests high likelihood of *have to* spreading to new environments, namely where subjective necessity is expressed in formal speech settings. More specifically, this extension towards new contexts could be a sign of ongoing grammaticalization in the system, as one of the parameters of grammatical change refers to the context-induced reinterpretation of new meanings of linguistic expressions (Heine 2003: 579; 2010: 405). Furthermore, even if the semi-modal *have to* is currently the default marker of root necessity in most or even all the varieties it shows varying underlying restructuring of constraints which are regionally specific and which correspond to some extent with previous accounts (Tagliamonte 2004; Tagliamonte & D'Arcy 2007; Tagliamonte & Smith 2006).

Overall, it can be said that the logistic regression method has proven useful and further analysis of variables, for example inter-speaker variation (e.g. age), could extend the scope of this statistical model to include sociolinguistic factors. To conclude, while the multivariate analysis essentially provides a microscopic configuration of estimates I suggest that the qualitative approach in addition to the variatinist method enhances the descriptive adequacy of the present study.

Finally, it may also be instructive to investigate *(have) got to* further and extend the analysis to other obligation/necessity markers such as *need to*, which has recently increased its use not only in standard BrE and AmE but also, as we have seen, in IrE, and particularly in JamE.

8 Conclusion

> *The facts of language change (and of language use) are as a rule subtler than the abstractions of linguists. Generalizations are no doubt part of what the art of linguistics is about, but empirical confrontations remain necessary if we wish to understand our own generalizations properly.* (Goossens 2000: 167)

The present study has focussed on variation and change associated with four modality markers (e.g. *must*, *have to*, *(have) got to* and *need t*o), which share the semantic domain of strong obligation/necessity in synchronic data from three New Englishes. The aim has been to trace the different paths of change in their use, and to find out whether these developments follow British, North American norms or independent 'local' ones in three specific sociolinguistic constellations:

- IrE, an old 'language-shift' variety in Europe, but a new L1 variety;
- JamE, a younger creole-based language-shift variety;
- IndE, a classic ESL.

Central to the study was the distinction between root and epistemic for these markers. Given that semi-modals are considered to mark informal style, register and stylistic effects were also noted. Since modality is one of the grammatical areas displaying the most fluctuations across written and spoken registers of present-day English (see Leech 2003; Mair & Leech 2006), it seemed necessary to uncover the potential distinctions/nuances of the modal – semi-modal alternation in the apparent time framework of other native and non-native varieties. The alternations between the modal *must* and semi-modal *have to* were assumed to be the interface for gradual change in different regional settings. To capture the relationship between the New Englishes and the two reference varieties, I adopted a further classification of an inner (IrE, BrE and AmE) and an outer circle (JamE and IndE).

By taking a multidimensional perspective on variation and change the following conclusion emerges from this study: the analysis provides evidence for intensified

language dynamics in JamE and IndE which is different from the three native varieties, IrE, BrE and AmE. In the course of the three analytical chapters (Chapters 5 to 7), it was possible to show that the distinction between the outer and inner circle is not solely restricted to diverging frequency patterns. Most significantly, the results confirm JamE and IndE to not only be different from native varieties in using such expressions, but also from each other, whereas IrE patterns are more similar to BrE. Furthermore, although the two outer circle varieties are not native-like, these findings do not support an interpretation from the perspective of 'colonial lag' either. This, in turn, seems to confirm the current status of JamE and IndE as emerging standards. The study has also confirmed that AmE, as found in data from SBC, is currently the most advanced variety in the use of strong obligation/necessity.

One major concern of this study has been to examine not only the frequency distribution of these markers but also their function and meaning in discourse context. Of particular interest was the question whether these varieties represent different diachronic stages in a pan-English trajectory of development of strong obligation/necessity. Several parameters have been considered to influence the synchronic profile of these items: morphosyntactic features, semantic contrast, discourse-pragmatic factors. To this end, the form-function distinction was correlated with frequency patterns in discourse.

The potential factors considered to be especially responsible for the increase of root *have to,* as opposed to informal *(have) got to* have been hypothesized to be: a) substrate influence in JamE (JC *hafi*); b) Americanization/colloquialization through the increase of informal speech habits; or c) further grammaticalization affecting the subjective/discourse-internal use of root *have to* as contrasting with objective/discourse-external (Depraetere & Verhulst 2008) use of root *must* or *need to*. In a nutshell, the behaviour of these modals was examined while taking convergence as participation in shared language-internal grammaticalization dynamics or/and divergence as displayed by possible substrate features, exhonormative influences or local socio-pragmatic stylistic conditions into account. From this perspective, the present study extends the research focus in this area as found, for example, in Collins (2009a; 2009b).

While the current study cannot offer an answer to the diachronic development of obligation/necessity in these three New Englishes, it identifies some synchronic

tendencies. In the following, I shall summarise the most important findings of the corpus-analysis related to the main issue of whether a reorganisation of strong obligation/necessity in the three varieties can be identified.

In terms of discourse frequency (Chapter 5), *have to* is the default marker of necessity in all the varieties under study. Modal *must* is most frequent in ICE-JA and ICE-India and shows almost equal distribution in both formal and informal conversations, as opposed to the other three varieties, where this marker displays a low incidence. By contrast, *(have) got to* is underrepresented but occurs more often in ICE-Ireland. In regards to *need to*, this marker displays surprisingly high discourse frequency in ICE-JA. The fact that *have to* is the most preferred marker of necessity, especially in JamE and IndE has been linked to an ongoing process of extension to cover more contexts, such as formal text types. Nevertheless, despite the current trends in present-day English regarding the spread of semi-modals in spoken language, the high incidence of *have to* in these two varieties cannot be explained by innovation either. On the other hand, the low frequency of informal *(have) got to* might be an indication that this recent Briticism has not yet become conventionalized in these varieties.

Further aspects addressed were, such as syntactic motivation, negation patterns, and stylistic patterns. In terms of morphosyntactic features the data from New Englishes confirms the recent trends found in BrE and AmE on the decrease of the negated *must* as compared to periphrastic *do*-negations with *have to*.

In Chapter 6, semantic contrast according to the root or epistemic meaning of these items was first related to discourse-frequency. Despite the apparent differences in distribution it turns out that such an approach does not capture the whole range of modal realizations. It was assumed that these markers are semantically interchangeable and that they can be systematically differentiated on the ground of the subjective – objective dichotomy. In many previous accounts on modality 'subjectivity', (i.e. the speaker's point of view or commitment to an action or proposition), is a feature of modals, whereas 'objectivity', (i.e. an external motivating force of necessity), is one of semi-modals. In contrast to the quantitative analysis which suggested that IndE and JamE share many similarities, the qualitative analysis in Chapter 6 has revealed different preferences in the use of these four markers in the two varieties.

However, even though the high frequency of modal *must* in ICE-JA and ICE-India might be seen as a conservative feature, the semantic analysis shows that most of the time it is used as it would be in BrE. A remarkable finding is that this modal is not used to express prototypical obligation in the sense of speaker-imposed obligation (see section 6.3.3.1). This seems to contrast previous findings on ICE-India in Collins (2009b), but confirms, to some extent, the receding trend in the use of this modal in present-day English. Another comparison can be made with the situation described in Biewer (2009) on South Pacific Englishes.

Subjectification (Traugott & Dasher 2002), as the cognitive process in semantic change, appears to be a valid theoretical basis to explain pragmatically motivated strategies of speaker-imposed obligation expressed with *have to* in the New Englishes data. Thanks to the diachronic evidence on the subjectification of obligation meaning for modals, it was possible to trace several intermediate steps of ongoing change within root necessity of this item in the ICE data. Undoubtedly a quantitative account alone would not have allowed for such insights.

A potential case of subjectification as conventionalization of implicature has been discussed in utterances which contain hedges (*I think*, *I believe*) + *have to* (see section 6.3.3.2). More important for the present case is that such utterances containing hedges seem to extend to other contexts which also cover formal conversations. It is, however, less clear-cut as to what extent the increased usage of subjective *have to* and *need to* is a sign of semantic change, as it might also involve language-external factors in the sense that speakers are aware of the less face-threatening effect of these semi-modals in present-day English. Based on such reasoning, the uses of the semi-modals *have to* and *need to* as found in ICE-JA can be interpreted as the result of the process of democratization in discourse. However, in the absence of a similar database for AmE, it was difficult to identify further relationships between these varieties. In addition, it was suggested that the similarity in meaning between *must* and *(have) got to* might be a factor which slows down the process of grammaticalization of this semi-modal in JamE and IndE.

In Chapter 7, ongoing change was presented as arising out of the interplay between morphosyntactic, discourse-pragmatic and semantic factors (root vs. epistemic, subjective vs. objective modality). The multivariate analysis has provided further

evidence for the current trends in the use of root readings of modal *must* and semi-modal *have to*. Several factors were identified to condition the choice of either *must* or *have to*. As it turns out, the semi-modal *have to* does not necessarily replace the modal, but is used generally to express a wide range of modal realizations, whereas *must* occurs in specialized contexts.

The other question of interest in this study was: what is the role of the local contact languages in modifying, speeding up or slowing down the pan-English grammaticalisation process? The lack of creole *hafi* in the Jamaican conversations, but also the low number of occurrences of informal *(have) got to* in the two outer circle varieties suggest that the two might follow a different norm than, for example, IrE. As stated/argued in Mair (2009b), standard educated JamE displays strong influences both from JC and North American English. Whether the preference for the semi-modal *have to* in JamE is a substrate influence and an American influence for *need to,* is a challenging question which should be addressed in future studies. In the case of IrE, there seems to be relatively strong evidence that the use of the semi-modal *have to* is associated with a non-standard vernacular use (see Joyce 1910; Tagliamonte & Smith 2006).

Putting these findings in the wider context of studies on the New Englishes, the marginalisation of certain items, e.g. *(have) got to* or *need to*, might suggest simplification in the system which was previously attributed to either learner strategies or to substrate influence (Biewer 2009: 51). On the other hand, a specific phenomenon such as the one discussed in Bao (2010) on English *must* in Singapore as "convergence-to-substratum" could not be identified in this stage of analysis. Further interesting details of potential similarities have been obtained by comparing this set of findings with those from non-standard British and Northern Irish varieties.

The analysis revealed a contrast between speech and writing, which suggests that differences in the use of modal necessity are more pronounced in spoken conversations. Overall, the high discourse frequency of the semi-modal *have to* correlates with the decreased formality of the spoken text categories. In addition, the various meanings of each marker seem to correlate with level of formality and register/text category. Also, the corpus data shows a contrast between speech and writing in the varying range of modal distribution. The analysis of the use of modal

necessity in written press texts has revealed that it conforms to the standard written norm. Moreover, these seem to be resistant to change, despite the fact that this genre has been shown to incorporate innovations more readily in present-day English (see Hundt & Mair 1999). This in turn, might suggest that the spoken language in these three varieties follows a different set of norms.

Apart from the descriptive approach on strong obligation/necessity, the present study aims to underline the relevance of both quantitative and qualitative approaches to the analysis of ongoing change within the modality system of the New Englishes. Although my analysis has offered insights on frequency patterns at several points, it has also shown that a more differentiated approach is required in order to capture ongoing change as a set of intermediate steps in a transition (see Goossens 2000).

The present study shows that ongoing change in the light of subjectification can be identified with qualitative analysis of the individual entries in the corpus. Nevertheless, a valid description of all the realizations in natural occurring data requires a set of solid criteria. Moreover, this does not discard the importance of quantitative evidence, only that it should be complemented, if possible, with more detailed qualitative analysis. In this sense, the discussion of root necessity in light of the subjective vs. objective dichotomy provides substantial evidence for the differences in the expression of obligation/necessity in these varieties.

Ultimately, univariate analyses remain a suitable method to describe the relationships between sets of data. By extension, the three-dimensional approach adopted in this study contributes to a multifaceted area of present-day English grammar such as modality.

On the one hand, the standardized corpus design currently available in *The International Corpus of English* has significantly contributed to the systematic comparison of regional and stylistic variation patterns. Specifically, private and public conversations provided the primary data source for the analysis. Given the usage-based approach, the parallel data sets incorporated in ICE have proven their usefulness. Moreover, it is a practical tool of comparison between native and non-native or emerging standard varieties.

On the other hand, the selected text categories also had some disadvantages. Although both the spoken conversations and the written press texts offered valuable

examples of natural occurring language, they sometimes turned out to be too small. Particularly disappointing was the fact that the Jamaican and Indian spoken dialogues hardly contained stylistic features signalling informality in discourse, such as contractions, e.g. '*ve to*, '*ve got to* or *gotta*. By contrast, the Irish conversations contained more occurrences of '*ve to*. Such findings raise several questions firstly about the role of the standard norm in JamE and IndE, and secondly whether more practical issues were involved, such as the criteria used in the transcription of recordings. A further point was that press texts were rather small in size which was also evident in the distribution of forms. Nevertheless, even if the database were larger, the coding of modal readings with each individual entry remains a task which should be performed by the researcher. This is an inevitable process given that gradience is a typical feature in modality.

As this study focused only on modality from a synchronic perspective, what still remains unsolved is the diachronic link between the three New Englishes and the two reference varieties, BrE and AmE. The recently edited collection of papers by Noël et al. (2014) is a significant step in this direction, as the contributions provide insight in the historical dimension of the modality system of postcolonial varieties in comparison with the parent variety, covering two native (AusE, WSAfE) and two non-native varieties (PhilE and BSAfE). While a uniform database of historical data of the New Englishes studied here is still unavailable, one must continue to rely on the evidence from secondary sources like BrE and AmE. It also might be useful to extend such a comparison to include other national varieties, e.g. AusE, NZE or CanE alongside these three New Englishes.

Lastly, it might be the case that an analysis of other markers of necessity, such as *should* and *ought to*, might change some of the assumptions from the present analysis. In this vein, it might be helpful to further investigate the meanings of these markers as well as their stylistic range. Especially *(have) got to* and *need to* will be interesting to analyse in the future in similar regionally stratified data as they revealed a rather marginal status in the present study. The present study, then, shows that the use of one modal item over the other in specific contexts in spoken data from New Englishes involves several stylistic and regional factors, which generate different ways of expressing modal obligation/necessity compared to other varieties of English.

Appendices

Appendix 1.a: Distribution of forms in spoken dialogues

(Semi-) modal	ICE-JA	ICE-India	ICE-Ireland	ICE-GB	SBC
Must	238 (66.1)	388 (107.7)	201 (55.8)	191 (53)	65 (26.1)
Have to	953 (264.7)	864 (240)	734 (203.8)	529 (146.9)	552 (221.6)
(Have) got to	16 (4.4)	15 (4.1)	57 (15.8)	183 (50.8)	119 (47.7)
Need to	250 (69.4)	38 (10.5)	107 (29.7)	122 (33.8)	122 (49.3)
Total*	1457 (404.7)	1305 (362.5)	1099 (305.2)	1025 (284.7)	858 (344.5)

Table 1.1. Distribution of *must*, *have to* and *(have) got to* and *need to* in spoken dialogues in JamE, IndE, IrE, BrE and AmE. The values in parentheses indicate the normalized frequencies per 100,000 words.
*Stands for all the morphological variants *have to*, *has to*, *having to*, *had to* as well as the contracted form *'ve to*; *have got to/has got to, got to* as well as contracted forms *'ve got to/'s got to*/*gotta*; *need to, needs to, needing to, needed to* including negated forms.

Appendix 1.b: Distribution of forms in written press texts

(Semi-)modal	ICE-JA	ICE-India	ICE-Ireland	ICE-GB	Frown
Must	22 (55)	11 (27.5)	29 (72.5)	11 (27.5)	31 (35.2)
Have to	40 (100)	28 (70)	42 (105)	48 (120)	65 (73.8)
(Have) got to	0	0	0	4 (10)	8 (9)
Need to	9 (22.5)	9 (22.5)	5 (12.5)	11 (27.5)	24 (27.2)
Total*	71 (177.5)	48 (120)	76 (190)	74 (185)	128 (145.4)

Table 1.2. Distribution of *must*, *have to* and *(have) got to* and *need to* in written press texts in JamE, IndE, IrE, BrE and AmE. The values in parentheses indicate the normalized frequencies per 100,000 words.
*Stands for all the morphological variants *have to*, *has to*, *having to*, *had to* as well as the contracted form *'ve to*; *have got to/has got to, got to* as well as contracted forms *'ve got to/'s got to*/*gotta*; *need to, needs to, needing to, needed to*.

Appendix 2.a: Semantic contrast – spoken texts

Must	**ICE-JA**	**ICE-IND**	**ICE-Ireland**	**ICE-GB**	**SBC**
Root	122 (33.8)	190 (52.7)	64 (17.7)	61 (16.9)	14 (5.6)
epistemic	34 (9.4)	125 (34.7)	101 (28)	99 (27.5)	49 (19.6)
Total	156 (43.3)	315 (87.5)	165 (45.8)	160 (44.6)	63 (25.3)

Table 2.1. Distribution of root and epistemic meanings with *must* in the spoken dialogues. The values in parentheses indicate the normalized frequencies per 100,000 words.

Have to	**ICE-JA**	**ICE-IND**	**ICE-Ireland**	**ICE-GB**	**SBC**
root	438 (121.6)	490 (136.1)	288 (80)	169 (46.9)	205 (82.3)
epistemic	42 (11.6)	44 (12.2)	25 (6.9)	20 (5.5)	16 (6.4)
Total	480 (133.3)	534 (148.3)	313 (86.9)	204 (56.6)	221 (88.7)

Table 2.2. Distribution of root and epistemic meanings with *have to* in the spoken dialogues. The values in parentheses indicate the normalized frequencies per 100,000 words.

(Have) got to	**ICE-JA**	**ICE-IND**	**ICE-Ireland**	**ICE-GB**	**SBC**
Root	8 (2.2)	12 (3.3)	48 (13.3)	150 (41.6)	90 (36.1)
Epistemic	5 (1.3)	3 (0.8)	3 (0.8)	12 (3.3)	9 (3.6)
Total	13 (3.6)	15 (4.1)	51 (14.1)	162 (45)	99 (39.7)

Table 2.3. Distribution of root and epistemic meanings with *(have) got to* in the spoken dialogues. The values in parentheses indicate the normalized frequencies per 100,000 words.

Need to	**ICE-JA**	**ICE-IND**	**ICE-Ireland**	**ICE-GB**	**SBC**
Root	162 (45)	21 (5.8)	53 (14.1)	72 (20)	68 (27.3)
Epistemic	9 (2.5)	2 (0.5)	5 (1.3)	3 (0.8)	3 (0.8)
Total	171 (47.5)	23 (6.3)	58 (16.1)	75 (20.8)	71 (19.7)

Table 2.4. Distribution of root and epistemic meanings with *need to* in the spoken dialogues. The values in parentheses indicate the normalized frequencies per 100,000 words.

Appendix 2.b: Semantic contrast – written press texts

Must	**ICE-JA**	**ICE-IND**	**ICE-Ireland**	**ICE-GB**	**Frown**
Root	17 (42.5)	5 (12.5)	19 (47.5)	8 (20)	22 (25)
Epistemic	4 (10)	2 (5)	6 (15)	3 (7.5)	7 (7.9)
Total	21 (52.5)	7 (17.5)	25 (62.5)	11 (27.5)	29 (32.9)

Table 2.5. Distribution of root and epistemic meanings with *must* in press texts. The values in parentheses indicate the normalized frequencies per 100,000 words.

Have to	**ICE-JA**	**ICE-IND**	**ICE-Ireland**	**ICE-GB**	**Frown**
Root	16 (40)	6 (15)	5 (12,5)	8 (20)	25 (28.4)
Epistemic	0	0	0	1	4 (4.5)
Total	16	6	5	9	29 (32.9)

Table 2.6. Distribution of root and epistemic meanings with *have to* in press texts. The values in brackets indicate the normalized frequencies per 100,000 words.

(Have) got to	**ICE-JA**	**ICE-IND**	**ICE-Ireland**	**ICE-GB**	**Frown**
Root	9 (22.5)	12 (30)	49 (122.5)	149 (372.5)	90 (102.2)
Epistemic	5 (1.3)	3 (0.8)	3 (0.8)	13 (3.6)	9 (2.5)
Total	14 (3.8)	15 (4.1)	52 (14.)	162 (45)	99 (27.5)

Table 2.7. Distribution of root and epistemic meanings with *(have) got to* in press texts. The values in parentheses indicate the normalized frequencies per 100,000 words.

Need to	**ICE-JA**	**ICE-IND**	**ICE-Ireland**	**ICE-GB**	**Frown**
Root	3 (7.5)	4 (10)	0	6 (15)	16 (18.8)
Epistemic	0	0	0	0	0
Total	3 (7.5)	4 (10)	0	6 (15)	16 (18.8)

Table 2.8. Distribution of root and epistemic meanings with *need to* in press texts. The values in parentheses indicate the normalized frequencies per 100,000 words.

Appendix 2.c: Distribution of sources within root necessity - spoken texts

	Discourse-external				
	ICE-JA	**ICE-India**	**ICE-Ireland**	**ICE-GB**	**Total**
must	71	119	23	30	**243**
	(35/36)	(34/85)	(4/19)	(10/20)	
have to	263	370	191	92	**916**
	(171/92)	(272/98)	(115/76)	(51/41)	
(have) got to	3	8	30	105	**146**
	(2/1)	(5/3)	(9/21)	(82/23)	
need to	75	8	20	39	**142**
	(46/27)	(4/4)	(12/8)	(14/24)	
Total	**414**	**505**	**264**	**266**	**1449**

Table 2.9. Distribution of raw numbers with discourse-external sources of necessity. The numbers in parentheses represent the distribution of sources in order of appearance in private and public dialogues, respectively.

	Discourse-internal				
	ICE-JA	**ICE-India**	**ICE-Ireland**	**ICE-GB**	**Total**
must	52	71	41	31	**195**
	(30/22)	(27/45)	(16/25)	(16/15)	
have to	175	120	97	77	**469**
	(104/71)	(71/49)	(48/49)	(33/44)	
(have) got to	5	4	18	45	**72**
	(4/1)	(2/2)	(4/14)	(25/20)	
need to	85	13	33	33	**164**
	(39/46)	(2/11)	(9/24)	(9/24)	
Total	**317**	**209**	**189**	**186**	**901**

Table 2.10. Distribution of raw numbers with discourse-internal sources of necessity. The numbers in parentheses represent the distribution of sources in order of appearance in private and public dialogues, respectively.

References

Aitchison, Jean. 2001[1991]. *Language Change: Progress or Decay?* Cambridge: Cambridge University Press.

Allsopp, Richard. 1996. In *Dictionary of Caribbean English Usage*. Oxford: Oxford University Press.

Baayen, R. Harald. 2008. *Analyzing Linguistic Data. A Practical Introduction to Statistics using R*. Cambridge Cambridge University Press.

Bailey, Beryl L. 1966. *Jamaican Creole Syntax*. Cambridge: Cambridge University Press.

Bailey, Beryl L. (ed.) 1971. *Jamaican Creole: Can dialect boundaries be defined?* (Pidginization and Creolization of Languages. Cambridge: Cambridge University Press.

Bailey, Guy. 2002. Real and Apparent Time. *The Handbook of Language Variation and Change*, ed. by Chambers, J.K., Peter Trudgill & Natalie Schilling-Estes, 312-332. Malden: Wiley-Blackwell Publishing.

Baker, Paul. 2010. *Sociolinguistics and Corpus Linguistics*. Edinburgh: Edinburgh University Press.

Bao, Zhiming. 2010. *Must* in Singapore English. *Lingua* 120.1727-1737.

Bauer, Laurie. 2002. *An Introduction to International Varieties of English*. Edinburgh: Edinburgh University Press.

Bayley, Robert. 2002. The Quantitative Paradigm. *The Handbook of Language Variation and Change*, ed. by Chambers, J.K., Peter Trudgill & Natalie Schilling-Estes, 117-141. Malden: Blackwell Publishing.

Bhatt, Rakesh M. 2004. Indian English: syntax. *A Handbook of varieties of English. Volume 2: Morphology and Syntax*, ed. by Kortmann, Bernd, Kate Burridge, Rajend Mesthrie, Edgar W. Schneider & Clive Upton, 1016-1030. Berlin: Mouton de Gruyter.

Biber, Douglas. 1988. *Variation across Speech and Writing*. Cambridge: Cambridge University Press.

Biber, Douglas. 1994. An Analytical Framework for Register Studies. *Sociolinguistic Perspectives on Register*, ed. by Biber, Douglas & Edward Finegan, 31-56. New York and Oxford: Oxford University Press.

Biber, Douglas. 2004. Modal use across registers and time. *Studies in the History of the English Language II: Unfolding Conversations*, ed. by Curzan, Anne & Kimberly Emmons, 189-216. Berlin/New York: Mouton de Gruyter.

Biber, Douglas. 2010. Corpus-Based and Corpus-Driven Analysis of Language Variation and Use. *The Oxford Handbook of Linguistic Analysis*, ed. by Heine, Bernd & Heiko Narrog, 159-191. Oxford: Oxford University Press.

Biber, Douglas & Susan Conrad. 2009. *Register, Genre, and Style*. Cambridge: Cambridge University Press.

Biber, Douglas, Susan Conrad & Randi Reppen. 1998. *Corpus Linguistics. Investigating Language Structure and Use*. Cambridge: Cambridge University Press.

Biber, Douglas & Edward Finegan. 1989. Drift and the evolution of English style: A history of three genres. *Language* 65.487-517.

Biber, Douglas, Stig Johansson, Geoffrey Leech, Susan Conrad & Edward Finegan. 1999. *Longman Grammar of Spoken and Written English*. London: Longman.

Bickerton, Derek. 1975. *Dynamics of a Creole System*. London/New York: Cambridge University Press

Biewer, Carolin. 2009. Modals and Semi-Modals of Obligation and Necessity in South Pacific Englishes. *Anglistik* 20.41-55.

Bod, Rens. 2010. Probabilistic Linguistics. *The Oxford Handbook of Linguistic Analysis*, ed. by Heine, Bernd & Heiko Narrog, 633-662. Oxford: Oxford University Press.

Brinton, Laurel J. 1991. The Origin and Development of Quasimodal *HAVE TO* in English. Paper presented at the Workshop on "The Origin and Development of Verbal Periphrases", 10th International Conference on Historical Linguistics (ICHL 10), August 16, 1991 University of British Columbia.1-31, Unpublished paper (retrieved on 07.01.2008 from http://faculty.arts.ubc.ca/lbrinton/HAVETO.PDF).

Bybee, Joan. 2003. Mechanisms of change in grammaticalization: The role of frequency. *Handbook of Historical Linguistics*, ed. by Joseph, Brian D. & Richard D. Janda, 602-623. Oxford: Blackwell Publishing.

Bybee, Joan (ed.) 2007. *Frequency of Use and the Organization of Language*. Oxford: Oxford University Press.

Bybee, Joan & Suzanne Fleischman (eds.) 1995. *Modality in grammar and discourse*. Philadelphia: John Benjamins Publishing.

Bybee, Joan & Paul Hopper (eds.) 2001. *Frequency and the Emergence of Linguistic Structure* (Typological Studies in Language (TSL) **45**). Amsterdam/Philadelphia: Benjamins.

Bybee, Joan, Michael R. Perkins & William Pagliuca. 1994. *The Evolution of Grammar: Tense, aspect, and modality in the languages of the world*. Chicago: University of Chicago Press.

Bybee, Joan & William Pagliuca. 1985. Crosslinguistic Comparison and the Development of Grammatical Meaning. *Historical Semantics/Historical Word Formation*, ed. by Fisiak, J., 59-84. Berlin: Mouton de Gruyter.

Campbell, Lyle & Richard Janda. 2001. Introduction: conceptions of grammaticalization and their problems. *Language Sciences* 23.93-112.

Christie, Pauline. 1989. Questions of standards and intra-regional differences in Caribbean examinations. *English across Cultures - Cultures across English: A Reader in Cross-Cultural Communication*, ed. by Garcia, Ofelia & Ricardo Otheguy, 243-262. Berlin: Mouton de Gruyter.

Christie, Pauline. 1991. Modality in Jamaican Creole. *Verb Phrase Patterns in Black English and Creole*, ed. by Edwards, Walter F. & Donald Winford, 223-239. Detroit: Wayne State University Press.

Christie, Pauline. 2003. *Language in Jamaica*. Kingston: Arawak Publications.

Coates, Jennifer. 1983. *The Semantics of the Modal Auxiliaries*. London and Canberra: Croom Helm.

Collins, Peter. 2014. Comparisons across British and American English Quasi-modals and Modals in Australian English Fiction 1800-1999, with Comparisons across British and American English. *Journal of English Linguistics* 42.7:30.

Collins, Peter C. 2005. The modals and quasi-modals of obligation and necessity in Australian English and other Englishes. *English World-Wide* 26.249-273.

Collins, Peter C. 2007. Can/could and may/might in British, American and Australian English: a corpus-based account. *World Englishes* 26.474-491.

Collins, Peter C. 2008. The English modals and semi-modals: Regional and stylistic variation. *The Dynamics of Linguistic Variation. Corpus evidence on English past and present*, ed. by Nevalainen, Terttu, Irma Taavitsainen, Päivi Pahta & Minna Korhonen, 129-145. Amsterdam/Philadelphia: John Benjamins.

Collins, Peter C. 2009a. *Modals and quasi-modals in English*. Amsterdam: Rodopi.

Collins, Peter C. 2009b. Modals and quasi-modals in world Englishes. *World Englishes* 28.281-292.

Collins, Peter, Ariane Macalinga Borlongan & Xinyue Yao Peter. 2014. Modality in Philippine English: A Diachronic Study. *Journal of English Linguistics* 42.68-88.

Corrigan, Karen P. 2000. 'What Bees To Be Maun Be'. Aspects of Deontic and Epistemic Modality in a Northern Dialect of Irish English. *English World-Wide* 21.25-62.

Croft, William. 2000. *Explaining Language change: An Evolutionary Approach*. London: Longman.

Cuyckens, Hubert, Kristin Davids & Lieven Vandelanotte. 2010. Introduction. *Subjectification, Intersubjectification and Grammaticalization*, ed. by Davidse, Kristin, Lieven Vandelanotte & Hubert Cuyckens, 1-26. Berlin/New Yoek: Mouton de Gruyter.

D'Souza, Jean. 1997. Indian English: some myths, some realities. *English World-Wide* 18.91-105.

DeCamp, David. 1971. Towards a generative analysis of a post-creole speech continuum. *Pidginization and Creolization of Languages*, ed. by Hymes, Dell, 349-370. Cambridge: Cambridge University Press.

Denison, David. 1993. *English Historical Syntax*. London/New York: Longman.

Depraetere, Ilse & An Verhulst. 2008. Source of modality: a reassessment. *English Language and Linguistics* 12.1-25.

Deuber, Dagmar. 2009. 'The English we speaking': Morphological and syntactic variation in educated Jamaican speech *Journal of Pidgin and Creole Languages* 24.1-52.

Deuber, Dagmar. 2010. Modal Verb Usage at the Interface of English and a Related Creole: A Corpus-based Study of *Can*/*Could* and *Will*/*Would* in Trinidadian English. *Journal of English Linguistics* 38.105-142.

Diaconu, Gabriela. 2012. Assessing Subjectivity and Objectivity in Modal Expressions from New Englishes: A Multivariate Analysis. *New trends and methodologies in applied English research II: Studies in variation, meaning and learning* ed. by Tizón-Couto, David, Bea Tizón-Couto, Iria Pastor- Gómez & Paula Rodriguez Puente, 37-72. Bern/Berlin etc.: Peter Lang.

Diller, Hans-Jürgen. 1999. Some thoughts on the stylistic function of contractions in written texts. *Form, function and variation in English*, ed. by U. Carls & P. Lucko, 235-245. Frankfurt: Peter Lang.

Dollinger, Stefan. 2006. The Modal Auxiliaries *have to* and *must* in the *Corpus of Early Ontario English*: Gradient Change and Colonial Lag. *Canadian Journal of Linguistics* 51.287-308.

Dollinger, Stefan. 2008. *New Dialect Formation in Canada*. Amsterdam and Philadelphia: John Benjamins Publishing Company.

Dorgeloh, Heidrun & Anja Wanner (eds.) 2010. *Syntactic Variation and Genre* (Topics in English Linguistics **70**). Berlin/New York: De Gruyter Mouton.

Duffley, Patrick J. 1994. *Need* and *dare*: The black sheep of the modal family. *Lingua* 94.213-243.
Esser, Jürgen. 2000. Medium-transferability and presentation structure in speech and writing. *Journal of Pragmatics* 32.1523-1538.
Facchinetti, Roberta, Manfred Kurg & Frank Palmer (eds.) 2003. *Modality in Contemporary English*. New York: Mouton de Gruyter.
Filppula, Markku. 1990. Substratum, Superstratum, and Universals in the Genesis of Hiberno-English. *Irish University Review. A Journal of Irish Studies: The English of the Irish: Special Issue* 20 41-54.
Filppula, Markku. 1991. Urban and rural varieties of Hiberno-English. *English around the world*, ed. by Cheshire, Jenny 51-60. Cambridge: Cambridge University Press.
Filppula, Markku. 1993. Changing Paradigms in the Study of Hiberno-English. *Irish University Review. A Journal of Irish Studies* 23 202-223.
Filppula, Markku. 1994. From Anglo-Irish to Hiberno-English. Paper presented to the Proceedings of the International Congress of Dialectologists, Bamberg 29.7.-4.8. 1990. Volume 3, Stuttgart, 1994.
Filppula, Markku. 1999. *The Grammar of Irish English*. London and New York: Routledge.
Filppula, Markku, Juhani Klemola & Heli Paulasto. 2008. *English and Celtic in Contact*. London: Routlage.
Fischer, Olga. 1994. Development of Quasi-Auxiliaries in English and Changes in Word Order. *Neophilologus* 78.137 - 164.
Fischer, Olga. 2003. Principles of grammaticalization and linguistic reality. *Determinants of Grammatical Variation in English*, ed. by Rohdenburg, Günter & Britta Mondorf, 445-478. Berlin/New York: Mouton de Gruyter.
Fischer, Olga, Muriel Norde & Harry Perridon (eds.) 2004. *Up and down the Cline - The Nature of Grammaticalization* (Typological Studies in Language. Amsterdam/Philadelphia: John Benjamins Publishing Company.
Fischer, Olga & Anette Rosenbach. 2000. Introduction. *Pathways of Change: Grammaticalization in English*, ed. by Fischer, Olga, Anette Rosenbach & Dieter Stein, 1-37. Amsterdam/Philadelphia: Benjamins.
Geisler, Christer. 2008. Statistical reanalysis of corpus data. *ICAME Journal* 32.35-45.
Giltrow, Janet. 2010. Genre as difference: The sociality of linguistic variation. *Syntactic Variation and Genre*, ed. by Dorgeloh, Heidrun & Anja Wanner, 29-51. Berlin/New York: De Gruyter Mouton.
Goossens, Louis. 1982. On the Development of the Modals and of the Epistemic Function in English. *Papers from the Fifth International Conference on Historical Linguistics*, ed. by Ahlqvist, Anders, 74-84. Amsterdam: Benjamins.
Goossens, Louis. 1984. The interplay of syntax and semantics in the development of the English modals. *English historical linguistics: Studies in development*, ed. by Blake, N. F. & Charles Jones, 149-159. Sheffield: University of Sheffield Printing Unit.
Goossens, Louis. 1987. The Auxiliarization of the English Modals: A Functional Grammar View. *Historical Development of Auxiliaries*, ed. by Harris, Martin & Paolo Ramat, 111-143. Berlin/New York: Mouton de Gruyter.
Goossens, Louis. 2000. Patterns of meaning extension, "parallel chaining", subjectification, and modal shifts. *Metaphor and Metonymy at the Crossroads:*

A Cognitive Perspective, ed. by Barcelona, Antonio, 149-169. Berlin/New York: Mouton de Gruyter.
Görlach, Manfred. 1987. Colonial lag? The alleged conservative character of American English and other 'colonial' varieties. *English World-Wide* 8.41-60.
Görlach, Manfred. 1991. English as a world language - the state of the art. *Englishes: Studies in Varieties of English 1984-1988*, ed. by Görlach, Manfred, 10-35. Amsterdam/Philadelphia: John Benjamins.
Görlach, Manfred (ed.) 1995. *More Englishes. New Studies in Varieties of English 1988-1994*. Amsterdam/Philadelphia: Benjamins.
Gotti, Maurizio, Marina Dossena, Dury Richard, Roberta Facchinetti & Maria Lima. 2002. *Variation in Central Modals: A Repertoire of Forms and Types of Usage in Middle English and Early Modern English*. Bern/Berlin: Peter Lang.
Greenbaum, Sidney. 1991. The development of the International Corpus of English. *English Corpus Linguistics: Studies in Honour of Jan Svartvik*, ed. by Aijmer, Karin & Bengt Altenberg, 83-91. London and New York: Longman.
Greenbaum, Sidney. 1992. A new corpus of English: ICE. *Directions in Corpus Linguistics. Proceedings of the Nobel Symposium 82, Stockholm, August 4-8 1991*, ed. by Svartvik, Jan, 171-179. Berlin/New York: Mouton de Gruyter.
Greenbaum, Sidney. 1996. Introducing ICE. *Comparing English World-Wide: The International Corpus of English*, ed. by Greenbaum, Sidney 3-12. Oxford: Oxford University Press.
Gries, Stefan. 2008. *Statistik für Sprachwissenschaftler*. Göttingen: Vandenhoeck&Rupprecht.
Groefsema, Marjolein. 1995. *Can*, *may*, *must* and *should*: a relevance theoretic account. *Journal of Linguistics* 31.53-79.
Haiman, John. 1998. Ritualization and the Development of Language. *Perspectives on Grammaticalization*, ed. by Pagliuca, William, 3-28. Amsterdam/Philadelphia: John Benjamins.
Hanson, Kristin. 1987. On Subjectivity and the History of Epistemic Expressions in English. *Papers from the 23rd Annual Regional Meeting od the Chicago Linguistic Society*, ed. by Need, Barbara, Eric Schiller & Anna Bosch, 133-147. Chicago: Chicago Linguistic Society.
Harris, John. 1984. English in the north of Ireland. *Language in the British Isles*, ed. by Trudgill, Peter, 115-134. Cambridge: Cambridge University Press.
Harris, John. 1991. Ireland. *English around the world*, ed. by Cheshire, Jenny 37-50. Cambridge: Cambridge University Press.
Harris, John. 1993. The grammar of Irish English. *Real English: The Grammar of English Dialects in the British Isles*, ed. by Milroy, James & Lesley Milory, 139-186. London and New York: Longman.
Haspelmath, Martin. 1999. Why is grammaticalization irreversible? *Linguistics* 37.1043-1068.
Heine, Bernd. 1993. *Auxiliaries: Cognitive Forces and Grammaticalization*. New York and Oxford: Oxford University Press.
Heine, Bernd. 2003. Grammaticalization. *The Handbook of Historical Linguistics*, ed. by Joseph, Brian D. & Richard D. Janda, 575-601. Oxford: Blackwell Publishing.

Heine, Bernd. 2010. Grammaticalization and Linguistic Analysis. *The Oxford Handbook of Linguistic Analysis*, ed. by Heine, Bernd & Heiko Narrog, 401-423. Oxford: Oxford University Press.
Heine, Bernd & Heiko Narrog. 2010. Introduction. *The Oxford Handbook of Linguistic Analysis*, ed. by Heine, Bernd & Heiko Narrog, 1-25. Oxford: Oxford University Press.
Heine, Bernd, Ulrike Claudi & Friederike Hünnemeyer. 1991. *Grammaticalization: a conceptual framework*. Chicago: University of Chicago Press.
Hickey, Raymond. 2004a. Development and diffusion of Irish English. *Legacies of Colonial English: Studies in Transported Dialects*, ed. by Hickey, Raymond, 82-120. Cambridge: Cambridge University Press.
Hickey, Raymond. 2004b. Englishes in Asia and Africa: origin and structure. *Legacies of Colonial English*, ed. by Hickey, Raymond, 503-535. Cambridge: Cambridge University Press.
Hickey, Raymond. 2004c. South Asian Engishes. *Legacies of Colonial English*, ed. by Hickey, Raymond, 536-558. Cambridge: Cambridge University Press.
Hickey, Raymond. 2005. Irish English in the context of previous research. *The Pragmatics of Irish English*, ed. by Barron, Anne & Klaus P. Schneider, 17-43. Berlin: Mouton de Gruyter.
Hickey, Raymond. 2007. *Irish English. History and present-day forms*. Cambridge: Cambridge University Press.
Hickey, Raymond. 2009. Modal Verbs in English and Irish. *Language Contacts Meet English Dialects: Studies in Honour of Markku Filppula*, ed. by Penttilä, Esa & Heli Paulasto, 259-274. Newcastle-upon-Tyne: Cambridge Scholars Publishing.
Hickey, Raymond (ed.) 2004d. *Legacies of Colonial English: Studies in Transported Dialects*. Cambridge: Cambridge University Press.
Hinrichs, Lars. 2006. *Codeswitching on the web: English and Jamaican Creole in e-mail communication*. Amsterdam: John Benjamins.
Hinrichs, Lars & Benedikt Szmrecsanyi. 2007. Recent changes in function and frequency of standard English genitives. *English Language and Linguistics* 11.437-474.
Hoffmann, Sebastian. 2004. Are low-frequency complex prepositions grammaticalized? On the limits of corpus data - and the importance of intuition. *Corpus Approaches to Grammaticalization in English*, ed. by Lindquist, Hans & Christian Mair, 171-210. Amsterdam/Philadelphia: John Benjamins.
Holm, John. 1986. The spread of English in the Caribbean area. *Focus on the Caribbean*, ed. by Görlach, Manfred & John A. Holm, 1-22. Amsterdam/Philadelphia: Benjamins.
Holm, John. 1988. *Pidgins and Creoles. Volume I: Theory and Structure*. Cambridge: Cambridge University Press.
Holm, John. 1994. English in the Caribbean. *The Cambridge History of the English Language. Volume V. English in Britain and Overseas: Origins and Development*, ed. by Burchfield, Robert, 328-381. Cambridge: Cambridge University Press.
Holmes, Janet. 1996. The New Zealand Spoken Component of ICE: Some Methodological Challenges. *Comparing English Worldwide*, ed. by Greenbaum, Sidney, 163-181. Clarendon Press: Oxford.

Hopper, Paul J. 1991. On Some Principles of Grammaticalization. *Approaches to Grammaticalization. Focus on Theoretical and Methodological Issues*, ed. by Traugott, Elisabeth & Bernd Heine, 17-35. Amsterdam and Philadelphia: John Benjamins.

Hopper, Paul J. & Elizabeth C. Traugott. 2003. *Grammaticalization (Second Edition)*. Cambridge: Cambridge University Press.

Huddleston, Rodney & Geoffrey Pullum. 2002. *Cambridge grammar of the English language*. Cambridge: Cambridge University Press.

Hundt, Marianne. 1998. *New Zealand English grammar - fact or fiction? A corpus-based study in morphosyntactci variation* Amsterdam/Philadelphia: John Benjamins.

Hundt, Marianne. 2009. Colonial lag, colonial innovation or simply language change? *One Language, Two Grammars? Differences between British and American English*, ed. by Rhodenburg, Günter & Julia Schlüter, 13-37. Cambridge: Cambridge University Press.

Hundt, Marianne & Christian Mair. 1999. 'Agile' and 'Uptight' Genres: The Corpus-based Approach to Language Change in Progress. *International Journal of Corpus Linguistics* 4.221-242.

Jacobsson, Bengt. 1979. Modality and the Modals of Necessity *Must* and *Have to*. *English Studies* 60.296-312.

Jacobsson, Bengt. 1994. Recessive and Emergent Uses of Modal Auxiliaries in English. *English Studies* 75.166-182.

Jankowski, Bridget. 2004. A transatlantic perspective of variation and change in English deontic modality. *Toronto Working Papers in Linguistics* 23.85-113.

Jantos, Susanne. 2009. Agreement in educated Jamaican English. Freiburg: Albert-Ludwig University PhD.

Jespersen, Otto. 1924. *The Philosophy of Grammar*. London: Allen & Unwin.

Johansson, Stieg & Signe Oksefjell. 1996. Towards a unified account of the syntax and semantics of *GET*. *Using Corpora for Language Research. Studies in Honour of Geoffrey Leech*, ed. by Thomas, Jenny & Mick Short, 57-75. London and New York: Longman.

Johnson, Keith. 2008. *Quantitative Methods in Linguistics*. Malden and Oxford: Blackwell Publishing.

Joyce, P. W. 1910. *English as We Speak in Ireland (Third Edition)*. Dublin: Wolfhound Press.

Kachru, Braj B. (ed.) 1982. *The Other Tongue: English Across Cultures*. Urbana: University of Illinois Press.

Kachru, Braj B. 1983. *The Indianization of English : the English language in India*. Delhi: Oxford University Press.

Kachru, Braj B. 1986. *The Alchemy of English. The spread, functiones and models of non-native Englishes*. Oxford and New York: Pergamon Press.

Kachru, Braj B. 1988. The sacred cows of English. *English Today* 16.3-8.

Kachru, Braj B. 1994. English in South Asia. *The Cambridge History of the English Language. Volume V. English in Britain and Overseas: Origins and Development*, ed. by Burchfield, Robert, 497-553. Cambridge: Cambridge University Press.

Kachru, Yamuna. 2008. Hindi-Urdu-Hindustani. *Language in South Asia*, ed. by Kachru, Braj B., Yamuna Kachru & Shikaripur N. Sridhar, 81-102. Cambridge: Cambridge University Press.

Kallen, Jeffrey L. 1997. Irish English: Context and Contacts. *Focus on Ireland*, ed. by Kallen, Jeffrey, 1-33. Amsterdam/Philadelphia: John Benjamins.

Kallen, Jeffrey L. & John M. Kirk. 2008. *ICE-Ireland: A User's Guide. Documentation to accompany the Ireland Component of the International Corpus of English (ICE-Ireland)*. Belfast: Cló Ollscoil na Banríona.

Kennedy, Graeme. 2002. Variation in the distribution of modal verbs in the British National Corpus. *Using Corpora to Explore Linguistic Variation*, ed. by Reppen, Randi, Susan M. Fitzmaurice & Douglas Biber, 73-90. Amsterdam/Philadelphia: John Benjamins.

Kirk, John M. 1994. Using VARBRUL for Studying Modal Auxiliary Verbs? *Hermes, Journal of Linguistics* 13.71-85.

Kirk, John M. & Jeffrey L. Kallen. 2005. Irish Standard English: How Celticised? How Standardised? *The Celtic Englishes IV. The Interface between English and the Celtic Languages. Proceedings of the fourth Colloquium on the "Celtic Englishes" held at the University of Potsdam in Golm (Germany) from 22-24 September 2004*, ed. by Tristram, Hildegard L. C., 88-113. Potsdam: Universitätsverlag Potsdam.

Kirk, John M. & Jeffrey L. Kallen. 2005 Irish Standard English: How Celticised? How Standardised? Paper presented to the The Celtic Englishes IV. The Interface between English and the Celtic Languages. Proceedings of the fourth Colloquium on the "Celtic Englishes" held at the University of Potsdam in Golm (Germany) from 22-24 September 2004, Potsdam, 2005

Kjellmer, Göran. 1998. On Contraction in Modern English. *Studia Neophilologica* 69.155-186.

Klinge, Alex. 1993. The English modal auxiliaries: from lexical semantics to utterance interpretation. *Journal of Linguistics* 29.315-357.

Klinge, Alex & H. Müller. 2005. *Modality. Studies in Form and Function*. Equinox: London.

Koch, Peter & Wulf Oesterreicher. 1985. Sprache der Nähe - Sprache der Distanz: Mündlichkeit und Schriftlichkeit im Spannungsfeld von Sprachtheorie und Sprachgeschichte. *Romanistisches Jahrbuch* 36.15-43.

Kortmann, Bernd, Kate Buridge, Rajend Mesthrie, Edgar W. Schneider & Clive Upton (eds.) 2004. *A Handbook of Varieties of English. Volume 2: Morphology and Syntax***2**). Berlin/New York: Mouton de Gruyter.

Kortmann, Bernd & Edgar W. Schneider. 2004. General Introduction. *The Handbook of Varieties of English. Volume 2: Morphology and Syntax*, ed. by Kortmann, Bernd, Kate Burridge, Rajend Mesthrie, Edgar W. Schneider & Clive Upton, 1-22. Berlin/New York: Mouton de Gruyter.

Krug, Manfred. 1998a. Gotta - the tenth central modal in English? Social, stylistic and regional variation in the British National Corpus as evidence of ongoing grammaticalization. Paper presented to the The Major Varieties of English, Växjö, 1998a.

Krug, Manfred. 1998b. String frequency: A cognitive motivating factor in coalescence, language processing and linguistic change. *Journal of English Linguistics* 26.286-320.

Krug, Manfred. 2000. *Emerging English Modals: A Corpus-based Study of Grammaticalization*. Berlin and New York: Mouton de Gruyter.
Krug, Manfred. 2003. Frequency as a determinant in grammatical variation and change. *Determinants of Grammatical Variation in English*, ed. by Rhodenburg, Günter & Britta Mondorf, 7-67. Berlin/New York: Mouton de Gruyter.
Kuteva, Tania. 2001. *Auxiliation: An Enquiry into the Nature of Grammaticalization*. Oxford: Oxford University Press.
Kytö, Merja. 1991. *Variation and Diachrony, with Early American English in Focus: Studies on CAN/MAY and SHALL/WILL*. Frankfurt am Main: Peter Lang.
Labov, William. 1982. Building on Empirical Foundations. *Perspectives on Historical Linguistics*, ed. by Lehmann, Winfred P. & Yakov Malkiel, 17-92. Amsterdam/Philadelphia: John Benjamins.
Labov, William. 1994. *Principles of Linguistic Change. Vol. I: Internal Factors*. Oxford: Blackwell.
Lalla, Barbara & Jean D'Costa. 1990. *Language in Exile*. Tuscaloosa and London: The University of Alabama Press.
Lass, Roger. 2000. Remarks on (uni)directionality. *Pathways of Change: Grammaticalization in English*, ed. by Fischer, Olga, Anette Rosenbach & Dieter Stein, 207-227. Amsterdam/Philadelphia: Benjamins.
Leech, Geoffrey. 1987. *Meaning and the English Verb (Second Edition)*. London and New York: Longman.
Leech, Geoffrey. 2003. Modality on the move: The English modal auxiliaries 1961-1992. *Modality in Contemporary English*, ed. by al., Roberta Facchinetti et, 223-240. New York: Mouton de Gruyter.
Leech, Geoffrey, Marianne Hundt, Christian Mair & Nicholas Smith. 2009. *Change in Contemporary English*. Cambridge Cambridge University Press.
Lehmann, Christian. 1995 [1982]. *Thoughts on Grammaticalization*. München-Newcastle: Lincom Europa.
Leitner, Gerhard. 1991. The Kolhapur Corpus of Indian English - intravarietal description and/or intervarietal comparison. *English Computer Corpora. Selected Papers and Research Guide*, ed. by Johansson, Stig & A. Stenström, 215-231. Berlin: Mouton de Gruyter.
Leitner, Gerhard. 1992. International Corpus of English: Corpus design - problems and suggested solutions. *New Directions in English Language Corpora*, ed. by Leitner, Gerhard, 33-64. Berlin/New York: Mouton de Gruyter.
Leitner, Gerhard. 2013. Pluricentricity and multiple layers of English in different habitats. Paper presented at the International Conference on Language Variation and Change in Postcolonial Contexts, LVC2013, Salerno.
Lightfoot, David W. . 1979. *Principles of Diachronic Syntax*. Cambridge: Cambridge University Press.
Lindquist, Hans & Christian Mair. 2004. Introduction. *Corpus Approaches to Grammaticalization in English*, ed. by Lindquist, Hans & Christian Mair, IX-XIV. Amsterdam/Philadelphia: Benjamins.
Lorenz, David. 2014. Contractions of English semi-modals: the emancipating effect of frequency. Freiburg: Albert-Ludwig University. PhD-Thesis.
Loureiro-Porto, Lucía. 2009. *The Semantic Predecessors of Need in the History of English (c750-1710)*. Malden: Wiley-Blackwell Publishing.
Lyons, J. 1977. *Semantics, vol. 2*. Cambridge: Cambridge University Press.

Mac Mathúna, Séamus. 2005. Remarks on Standardisation in Irish English, Irish and Welsh. Paper presented to the The Interface between English and the Celtic Languages. Proceedings of the fourth Colloquium on the "Celtic Englishes" held at the University of Potsdam in Golm (Germany) from 22-24 September 2004, Potsdam, 2005.

Mair, Christian. 1992. Problems in the compilation of a corpus of Standard Caribbean English: A pilot study. *New Directions in English Language Corpora*, ed. by Leitner, Gerhard, 75-96. Berlin: Mouton de Gruyter.

Mair, Christian. 1997. Parallel Corpora: A Real-Time Approach to the Study of Language Change in Progress. *Corpus-based Studies in English. Papers from the seventeenth International Conference on English Language Research on Computerized Corpora (ICAME 17) Stockholm, May 15-19, 1996*, ed. by Ljung, Magnus, 195-209. Amsterdam and Atlanta: Rodopi.

Mair, Christian. 1998. Corpora and the study of the major varieties of English: Issues and results. Paper presented to the The Major Varieties of English. Papers from MAVEN 97, Växjö 20-22 November 1997, 1998.

Mair, Christian. 2002a. Creolisms in an emerging standard. Written English in Jamaica. *English World-Wide* 23 31-58.

Mair, Christian. 2002b. Three changing patterns of verb complementation in Late Modern English: a real-time study based on matching text corpora. *English Language and Linguistics* 6.105-131.

Mair, Christian. 2003. Kreolismen und verbales Identitätsmanagement im geschriebenen jamaikanischen English. *Zwischen Ausgrenzung und Hybridisierung*, ed. by Antonia Napp & Wolfram Lutterer, 79-96. Würzburg: Ergon.

Mair, Christian. 2004. Corpus linguistics and grammaticalisation theory. Statistics, frequencies, and beyond. *Corpus Approaches to Grammaticalization in English*, ed. by Lindquist, Hans, 123-150. Amsterdam: John Benjamins Publishing Company.

Mair, Christian. 2007. British English/American English grammar: Convergence in writing - divergence in speech? *Anglia* 125.84-100.

Mair, Christian. 2009a. Corpora and the study of recent change in language. *Corpus Linguistics: An International Handbook, Volume 2*, ed. by Lüdeling, Anke & Merja Kytö, 1109-1125. Berlin/NewYork: Walter de Gruyter.

Mair, Christian. 2009b. Corpus linguistics meets sociolinguistics: The role of evidence in the study of sociolinguistic variation and change. *Corpus Linguistics: Refinements and Reassessments*, ed. by Antoinette Renouf & Andrew Kehoe, 7-32. Amsterdam: Rodopi.

Mair, Christian & Geoffrey Leech. 2006. Current changes in English Syntax. *The Handbook of English Linguistics*, ed. by Aarts, Bas & April McMahon, 318-342. Malden, Oxford and Victoria: Wiley-Blackwell.

Mair, Christian & Andrea Sand. 1998. Caribbean English: Structure and Status of an Emerging Variety. *Anglistentag 1997 Gießen: Proceedings*, ed. by Borgmeier, Raimund, Herbert Grabes & Andreas H. Jucker, 187-198. Trier: Wissenschaftlicher Verlag Trier.

Marckwardt, Albert H. 1980. *American English*. New York?: Oxford University Press.

McArthur, Tom. 2003. World English, Euro-English, Nordic English. *English Today* 19 54-58.

McEnery, Tony, Richard Xiao & Yukio Tono. 2006. *Corpus-Based Language Studies: An advanced resource book*. London and New York: Routledge.
McEnery, Tony & Andrew Wilson. 2003 [1996]. *Corpus Linguistics*. Edinburgh: Edinburgh University Press.
McMahon, April. 1994. *Understanding language change*. Cambridge: Cambridge University Press.
Mehrotra, Raja Ram. 1998. *Indian English*. Amsterdam/Philadelphia: John Benjamins Publishing Company.
Mesthrie, Rajend. 2006. World Englishes and the multilingual history of English. *World Englishes* 25.381-390.
Mesthrie, Rajend & Bhatt Rakesh M. 2008. *World Englishes. The Study of New Linguistic Varieties*. Cambridge Cambridge University Press.
Meyer, Charles F. 2002. *English Corpus Linguistics: An introduction*. Cambridge: Cambridge University Press.
Millar, Neil. 2009. Modal verbs in TIME. Frequency changes 1923-2006. *International Journal of Corpus Linguistics* 14 191-220.
Miller, Jim. 1993. The Grammar of Scottish English. *Real English: The Grammar of English Dialects in the British Isles*, ed. by Milroy, James & Lesley Milory, 99-138. London and New York: Longman.
Miller, Jim. 2004. Spoken English in Scotland. *The Handbook of English Varieties. Volume 2: Morphology and Syntax*, ed. by Kortmann, Bernd, Kate Buridge, Rajend Mesthrie, Edgar W. Schneider & Clive Upton, ---. Berlin/New York: Moutin de Gruyter.
Miller, Jim & Keith Brown. 1982. Aspects of Scottish English Syntax. *English World-Wide* 3.3-17.
Milson-Whyte, Vivette 2006. A Comparison of the Structural Composition of the Auxiliary in Standard American English and in Jamaican Creole. *Arizona Working Papers in Second Language Acquisition and Teaching in SLAT* 13.93-116.
Mindt, Dieter. 1995. *An Empirical Grammar of the English Verb: Modal Verbs*. Berlin: Cornelsen.
Mufwene, Salikoko. 2001. *The Ecology of Language Evolution*. Cambridge: Cambridge University Press.
Mufwene, Salikoko. 2006. Pidgins and Creoles. *The Handbook of World Englishes*, ed. by Kachru, Yamuna & Cecil L. Nelson, 313-327. Malden Blackwell Publishing.
Mukherjee, Joybrato. 2007. Steady states in the evolution of new Englishes. Present-day Indian English as an equilibrium. *Journal of English Linguistics* 35.157-187.
Mukherjee, Joybrato & Stefan Th. Gries. 2009. Collostructional nativisation in New Englishes. Verb-construction associations in the International Corpus of English. *English World-Wide* 30 27-51.
Mukherjee, Joybrato & Sebastian Hoffmann. 2006. Describing verb-complememtational profiles of New Englishes: A pilot study of Indian English. *English World-Wide* 27.147-173.
Myhill, John. 1995. Change and continuity in the functions of the American English modals. *Linguistics* 33.157-211.
Myhill, John. 1996. The Development of the Strong Obligation System in American English. *American Speech* 71 339-388.

Nagle, Stephen J. 1989. *Inferential Change and Syntactic Modality in English*. Frankfurt am Main: Verlag Peter Lang.
Nagle, Stephen J. 2003. Double modals in the southern United States: Syntactic structure or syntactic structures? *Modality in Contemporary English*, ed. by Facchinetti, Roberta, Manfred Krug & Frank Palmer, 349-371. Berlin/New York: Mouton de Gruyter.
Narrog, Heiko. 2010. (Inter)subjectification in the domain of modality and mood - Concepts and cross-linguistic realities. *Subjectification, Intersubjectification and Grammaticalization*, ed. by Davidse, Kristin, Lieven Vandelanotte & Hubert Cuyckens, 385-429. Berlin/New York: Mouton de Gruyter.
Nelson, Gerald. 1996a. The Design of the Corpus. *Comparing English Worldwide*, ed. by Greenbaum, Sidney. Oxford: Clarendon Press.
Nelson, Gerald. 1996b. Markup Systems. *Comparing English Worldwide: The International Corpus of English* ed. by Greenbaum, Sidney, 36-53. Oxford: Clarendon Press.
Nelson, Gerald. 2006. The core and the periphery of world Englishes: a corpus-based exploration. *World-Englishes* 25.115-129.
Nelson, Gerald, Sean Wallis & Bas Aarts. 2002. *Exploring Natural Language: Working with the British Component of the International Corpus of English*. Amsterdam/Philadelphia: John Benjamins.
Newmeyer, Frederick J. 1998. *Language Form and Language Function*. Cambridge, MA: MIT Press.
Noël, Dirk, Bertus van Rooy & Johan van der Auwera (eds.). 2014. Special Issue: Diachronic Approaches to Modality in World Englishes. *Journal of English Linguistics* 42.3-88.
Nokkonen, Soili. 2006. The semantic variation of NEED TO in four recent British English corpora. *International Journal of Corpus Linguistics* 11.29-71.
Nordlinger, Rachel & Elizabeth C. Traugott. 1997. Scope and the development of epistemic modality: evidence from *ought to*. *English Language and Linguistics* 1.295-317.
Nuyts, Jan. 2001. *Epistemic Modality, Language and Conceptualization. A Cognitive-Pragmatic Perspective*. Amsterdam/Philadelphia: John Benjamins.
Oakes, Michael P. 1998. *Statistics for Corpus Linguistics*. Edinburgh: Edinburgh University Press.
Odlin, Terence. 1994. A demographic perspective on the shift from Irish to English. *The Germanic Mosaic: Cultural and Linguistic Diversity in Society*, ed. by Blackshire-Belay, C. A., 137-145. Westport, CT: Greenwood Press.
Palmer, Frank. 1990 [1979]. *Modality and the English Modals (Second Edition)*. London and New York: Longman Linguistics Library.
Palmer, Frank. 2001. *Mood and modality (Second Edition)*. Cambridge: Cambridge University Press.
Palmer, Frank. 2003. Theoretical, descriptive and typological issues. *Modality in Contemporary English*, ed. by Facchinetti, Roberta et al. . New York: Mouton de Gruyter.
Papafragou, Anna. 2000. *Modality: Issues in the Semantics-Pragmatics Interface*. Amsterdam and Lausanne: Elsevier.
Patrick, Peter L. 1999. *Urban Jamaican Creole: Variation in the Mesolect*. Amsterdam and Philadelphia: John Benjamins.

Patrick, Peter L. 2004. Jamaican Creole: morphology and syntax. *The Handbook of Varieties of English. Volume 2: Morphology and Syntax*, ed. by Kortmann, Bernd, Kate Burridge, Rajend Mesthrie, Edgar W. Schneider & Clive Upton, 407-438. Berlin/New York: Mouton de Gruyter.

Perkins, Michael R. 1983. *Modal Expressions in English*. London: Frances Pinter.

Plank, Frans. 1984. The modals story retold. *Studies in language* 8.305-364.

Platt, John T., Heidi Weber & Mian L. Ho. 1984. *The New Englishes*. London: Routledge.

Poplack, Shana & Sali Tagliamonte. 1996. Nothing in context: past temporal reference in Nigerian Pidgin English. *Changing meanings, changing functions: Papers relating to grammaticalization in contact languages*, ed. by Baker, Philip & Anand Syea, 71-94. Westminster: University of Westminster Press.

Poplack, Shana & Sali Tagliamonte. 1999. The grammaticization of *going to* in (African American) English. *Language Variation and Change*.315-342.

Pride, John B. (ed.) 1982. *New Englishes*. Rowley, London and Tokyo: Newbury House Publisher.

Quirk, Randolph, Sidney Greenbaum, Geoffrey Leech & Jan Svartvik. 1985. *A Comprehensive Grammar of the English Language*. London: Longman.

Rhodenburg, Günter & Julia Schlüter. 2009. New departures. *One Language, Two Grammars? Differences between British and American English*, ed. by Rhodenburg, Günter & Julia Schlüter, 364-423. Cambridge: Cambridge University Press.

Rickford, John R. 1987. *Dimensions of a Creole Continuum*. Stanford CA: Stanford University Press.

Rohdenburg, Günter. 1996. Cognitive complexity and increased grammatical explicitness in English. *Cognitive Linguistics* 7.149-182.

Ronan, Patricia. 2011. Modal *would* and *will* in ICE Ireland. Paper presented at the CHRONOS 10: 10th International Conference on Tense, Aspect, Modality and Evidentiality, Aston University, Birmingham.

Rossouw, Ronel & Bertus van Rooy. 2011. Modality on trek: Explaining diachronic changes in written South African English across text and context. Paper presented at the Chronos 10: 10th International Conference on Tense, Aspect, Modality and Evidentiality, Aston Unvirsity, Birmingham.

Sailaja, Pingali. 2009. *Indian English*. Edinburgh: Edinburgh University Press.

Salkie, Raphael, Pierre Busuttil & Johan van der Auwera (eds.) 2009. *Modality in English: Theory and Description*. Berlin/New York: Mouton de Gruyter.

Sand, Andrea. 1999. *Linguistic Variation in Jamaica - A Corpus-Based Study of Radio and Newspaper Usage*. Tübingen: Gunter Narr Verlag.

Sand, Andrea. 2005. Angloversals? Shared Morpho-Syntactic Features in Contact Varieties of English. Habilitationsschrift.

Sankoff, David. 1988. Sociolinguistics and syntactic variation. *Linguistics: The Cambridge Survey. Volume IV: Language; the socio-sultural context*, ed. by Newmeyer, Frederick J., 140-161. Cambridge: Cambridge University Press.

Sankoff, David & William Labov. 1979. On the uses of variable rules. *Language in Society* 8.189-222.

Schilling-Estes, Natalie. 2002. Investigating Stylistic Variation. *The Handbook of Language Variation and Change*, ed. by Chambers, J.K., Peter Trudgill & Natalie Schilling-Estes, 375-401. Malden: Blackwell Publishing.

Schmied, Josef. 1990. Corpus linguistics and non-native varieties of English. *World Englishes* 9.255-268.
Schmied, Josef. 1994. Syntactic style variation in Indian English. *Anglistentag 1993 Eichstätt: Proceedings*, ed. by Blaicher, Günther & Brigitte Glaser, 217-232. Tübingen: Niemayer.
Schmied, Josef. 1996. Second-Language Corpora. *Comparing English Worldwide: The International Corpus of English*, ed. by Greenbaum, Sidney, 182-195. Oxfrod: Clarendon.
Schneider, Edgar W. 2004. How to trace structural nativization: Particle verbs in world Englishes. *World Englishes* 23.227-281.
Schneider, Edgar W. . 2003. The dynamics of New Englishes: From identity construction to dialect birth. *Language* 79.233-281.
Schneider, Edgar W. . 2007. *Postcolonial English. Varieties around the world*. Cambridge Cambridge University Press.
Schulz, Monika Edith. 2012. Morphosyntactic Variation in British English Dialects: Evidence from Possession, Obligation and Past Habituality. Freiburg: Albert-Ludwig University. PhD-Thesis (retrieved on 18.06.2012 from http://www.freidok.uni-freiburg.de/volltexte/8511).
Sebba, Mark. 1993. *London Jamaican: Language Systems in Interaction*. London and New York: Longman.
Sedlatschek, Andreas. 2009. *Contemporary Indian English*. Amsterdam and Philadelphia: John Benjamins Publishing Company.
Shastri, S. V. 1988. The Kolhapur corpus of Indian Enlish and work done on its basis so far. *ICAME Journal* 12.
Shepherd, Susan. 1982. From deontic to epistemic: An analysis of modals in the history of English, creoles and language aquisition. *Papers from the Fifth International Conference on Historical Linguistics*, ed. by Ahlqvist, Anders, 316-323. Amsterdam: Benjamins.
Shields-Brodber, Kathryn. 1997. Requiem for English in an 'English-speaking' community. The case of Jamaica. *Englishes Around the World 2. Caribbean, Africa, Asia, Australasia. Studies in honour of Manfred Görlach*, ed. by Schneider, Edgar W. , 57-67. Amsterdam and Philadelphia: John Benjamins Publishing Company.
Shields, Kathryn. 1989. Standard English in Jamaica: A case of competing models. *English World-Wide* 10.41-53.
Siemund, Peter. 2005. Independent Developments in the Genesis of Irish English. Paper presented to the The Celtic Englishes IV. The Interface between English and the Celtic Languages. Proceedings of the fourth Colloquium on the "Celtic Englishes" held at the University of Potsdam in Golm (Germany) from 22-24 September 2004, Potsdam, 2005.
Smith, Nicholas 2003. Changes in the modals and semi-modals of strong obligation and epistemic necessity in recent British English. *Modality in Contemporary English*, ed. by Facchinetti, Roberta et al. , 241-266. New York: Mouton de Gruyter.
Sweetser, Eve E. 1982. Root and Epistemic Modals: Causality in Two Worlds. *Proceedings of the Eighth Annual Meeting of the Berkeley Linguistics Society, February 13-15, 1982*, ed. by Macaulay, Monica & Orin D. Gensler, 484-507. Berkeley: Berkeley Linguistics Society.

Sweetser, Eve E. 1990. *From Etymology to Pragmatics: Metaphorical and Cultural Aspects of Semantic Structure*. Cambridge: Cambridge University Press.

Szmrecsanyi, Benedikt. 2006. *Morphosyntactic Persistence in Spoken English*. Berlin and New York: Mouton de Gruyter.

Taeymans, Martine. 2004. An investigation into the marginal modals DARE and NEED in British present-day English: A corpus-based approach. *Up and down the Cline - The Nature of Grammaticalization*, ed. by Fischer, Olga, Muriel Norde & Harry Perridon, 97-114. Amsterdam/Philadelphia: Benjamins.

Tagliamonte, Sali. 2002. Comparative Sociolinguistics. *The Handbook of Language Variation and Change*, ed. by Chambers, J.K., Peter Trudgill & Natalie Schilling-Estes, 729-763. Malden: Blackwell Publishing.

Tagliamonte, Sali. 2004. 'Have to, gotta, must': Grammaticalisaton, variation, and specialisation. *Corpus Approaches to Grammaticalization in English*, ed. by Lindquist, Hans 33-56. Amsterdam: John Benjamins Publishing Company.

Tagliamonte, Sali & Alexandra D'Arcy. 2007. The modals of obligation/necessity in Canadian perspective. *English World-Wide* 28 47-87.

Tagliamonte, Sali & Jennifer Smith. 2006. Layering competition and a twist of fate. *Diachronica* 23.341-380.

Teubert, Wolfgang. 2005. My version of corpus linguistics. *International Journal of Corpus Linguistics* 10.1-13.

Tinkham, Thomas. 1993. Sociocultural variation in Indian English speech acts. *World Englishes* 12.239-247.

Traugott, Elisabeth. 1988. Pragmatic strengthening and grammaticalization. Paper presented to the Proceedings of the Forteenth Annual Meeting of the Berkeley Linguistics Society, February 13-15, 1988. General Session and Parasession on Grammaticalization, Berkeley, 1988.

Traugott, Elisabeth. 2003. From Subjectification to Intersubjectification. *Motives for Language Change*, ed. by Hickey, Raymond, 124-139. Cambridge: Cambridge University Press.

Traugott, Elisabeth Closs. 1997. Subjectification and the development of epistemic meaning: The case of *promise* and *threaten*. *Modality in Germanic Languages: Historical and Comparative Perspectives*, ed. by Swan, Toril & Olaf Jansen Westwik, 185-210. Mouton de Gruyter: Berlin/New York.

Traugott, Elizabeth Closs. 1989. On the rise of epistemic meanings in English: an example of subjectification in semantic change. *Language* 65.31 - 55.

Traugott, Elizabeth Closs. 1995. Subjectification in grammaticalization. *Subjectivity and Subjectivisation*, ed. by Stein, Dieter & Susan Wright, 31-54. Cambridge: Cambridge University Press.

Traugott, Elizabeth Closs. 1999. Why must is not moot. Paper presented at the Paper presented at the Fourteenth International Conference on Historical Linguistics, Vancouver, Canada.

Traugott, Elizabeth Closs. 2010. (Inter)subjectivity and (inter)subjectification: A reassessment. *Subjectification, Intersubjectification and Grammaticalization*, ed. by Davidse, Kristin, Lieven Vandelanotte & Hubert Cuyckens, 29-71. Berlin/NewYork: Mouton de Gruyter.

Traugott, Elizabeth Closs & Richard B. Dasher. 2002. *Regularity in Semantic Change*. Cambridge: Cambridge University Press.

Traugott, Elizabeth Closs & Bernd Heine. 1991. Introduction. *Approaches to Grammaticalization. Volume I: Focus on the Theoretical and Methodological Issues*, ed. by Traugott, Elizabeth C. & Bernd Heine, 1-14. Amsterdam/Philadelphia: John Benjamins.

Tregidgo, P.S. 1982. Must and may: demand and permission. *Lingua* 56.75-92.

Trousdale, Graeme. 2003. Simplification and redistribution. An account of modal verb usage in Tyneside English. *English World-Wide* 24.271-284.

Trudgill, Peter & Jean Hannah. 2002. *International English: A guide to varieties of English, 4th edition*. London: Arnold.

Tsangalidis, Anastasios & Roberta Facchinetti (eds.) 2009. *Studies on English modality. In Honor of Frank Palmer*. Bern/Berlin: Peter Lang.

van der Auwera, Johan & Vladimir A. Plungian. 1998. Modality's Semantic Map. *Linguistic Typology* 2.79-124.

van Rooy, Bertus & Ronel Rossouw. 2011. Modality in Black South African English. Paper presented at the Chronos 10: 10th International Conference on Tense, Aspect, Modality and Evidentiality, Aston University, Birmingham.

van Rooy, Bertus & Ronel Wasserman. 2014. Do the Modals of Black and White South African English Converge? *Journal of English Linguistics* 42.51-67.

Vendler, Zeno. 1967. *Linguistics in Philosophy*. Ithaca, New York: Cornell University Press.

Verstraete, Jean-Christophe. 2001. Subjective and objective modality: Interpersonal and ideational functions in the English modal auxiliary system. *Journal of Pragmatics* 33.1505-1528.

Visconti, Jacqueline. 2004. Conditionals and subjectification: Implications for a theory of semantic change. *Up and down the Cline - The Nature of Grammaticalization*, ed. by Fischer, Olga, Muriel Norde & Harry Peridon, 169-192. Amsterdam/Philadelphia: John Benjamins.

Visser, F. Th. 1969. *An Historical Syntax of the English Language. Part Three, First Half: Syntactical Units with Two Verbs*. Leiden: E.J. Brill.

Warner, Anthony. 1990. Reworking the history of English auxiliaries. *Papers from the 5th International Conference on English Historical Linguistics*, ed. by Adamson, Sylvia, Vivien Law, Nigel Vincent & Susan Wright. Amsterdam/Philadelphia: Benjamins.

Warner, Anthony. 1993. *English Auxiliaries: Structure and History*. Cambridge: Cambridge University Press.

Wärnsby, Anna. 2009. On Controlability as a Contextual Variable. *Studies on ENglish Modality: In Honour of Frank Palmer*, ed. by Tsangalidis, Anastasios & Roberta Facchinetti, 69-97. Berlin: Peter Lang.

Wasserman, Ronel & Bertus van Rooy. 2014. The Development of Modals of Obligation and Necessity in White South African English through Contact with Afrikaans. *Journal of English Linguistics* 42.31-50.

Westney, Paul. . 1995. *Modals and Periphrastics in English: An investigation into the semantic correspondence between certain English modal verbs and their periphrastic equivalents*. Tübingen: Max Niemeyer.

Wierzbicka, Anna. 1987. The semantics of modality. *Folia Linguistica* 21.25-43.

Wilson, Andrew. 2005. Modal verbs in written Indian English: A quantitative and comparative analysis of the Kolhapur corpus using correspondence analysis. *ICAME Journal* 29.151-169.

Winford, Donald. 1993. *Predication in Caribbean English Creoles*. Amsterdam: John Benjamins Publishing Company.

Zusammenfassung in deutscher Sprache

Die vorliegende Arbeit untersucht verbale Modalkonstruktionen der Pflicht und Notwendigkeit, z.B. *must*, *have to*, *(have) got to* und *need to* als regionale und stilistische Variationsmuster in den neuen Standardvarietäten des Englischen (auch „New Englishes“). Die Untersuchung schließt sich somit früheren Studien (Krug 2000; Leech 2003; Mair & Leech 2006; Smith 2003) an, die aktuell ablaufende Sprachwandelprozesse innerhalb des Modalsystems des Englischen beschrieben haben.

Während die genannten Studien auf empirischen Untersuchungen schriftlicher britischer und amerikanischer Texte basieren, werden in der vorliegenden Analyse gesprochene Texte dreier Standardvarietäten des Englischen, nämlich des jamaikanischen Englisch, des indischen Englisch und des irischen Englisch verwendet. Somit soll ein umfassenderes Bild über verschiedene Grammatikalisierungstendenzen auch in anderen ex-kolonialen Standardvarietäten des Englischen weltweit widergeben werden. In diesem Zusammenhang sind besonders das jamaikanische und das irische Englisch wenig untersucht worden.

Ziel dieser Arbeit ist es herauszufinden, ob diese Änderungen in den drei Varietäten zutreffen und ob diese unterschiedlichen Grammatikalisierungspfaden im Vergleich zum britischen und amerikanischen Englisch folgen. Diese Perspektive folgt jüngsten Forschungsentwicklungen im Bereich der neueren Standardvarietäten des Englischen, die zusätzlich zur traditionellen britischen Norm auch auf externe Einflüsse im sprachlichen Verhalten der Sprecher hindeuten (z.B. erhöhter Einsatz von typischen Formen der amerikanischen Umgangssprache, auch bekannt als *colloquialization*, die eine stilistische Bedeutung haben). Daraus ergibt sich die linguistische Anerkennung von neu entstehenden Strukturen als regionale und „lokale“ Normen.

Die Grammatikalisierungspfade werden zum einen im epistemischen und deontischen Kontrast (Coates 1983), zum anderen, spezifisch im deontischen Gebrauch als *subjektiv* oder *objektiv* (Depraetere & Verhulst 2008) hervorgehoben. Desweiteren wurde die Relevanz stilistischer Merkmale (formell vs. informell) von Modalkonstruktionen näher erläutert. Um diese Ziele systematisch erreichen zu können, werden qualitative und quantitative Forschungsansätze ergänzend eingesetzt. Als sprachliche Quellen habe ich gesprochene formelle und informelle Gespräche (180 Gespräche mit jeweils 2.000 Wörter im Textformat) der drei Varietäten innerhalb des *International Corpus of English* (ICE) untersucht, nämlich ICE-India, ICE-Jamaica und ICE-Ireland. Desweiteren vergleiche ich die drei obengenannten Datensätze mit der britischen Komponente ICE-GB und mit dem *Santa Barbara Corpus of Spoken American English*, das das fehlende ICE Korpus mit gesprochenen amerikanischen Texten ersetzt.

Die detailierte Korpus-Analyse ist von der Hypothese ausgegangen, dass es einen Unterschied gibt zwischen den Varietäten die dem *outer circle* bzw. dem *inner circle* (Kachru 1988) angehören. Das bedeutet, dass der Gebrauch der Modalkontruktionen der Pflicht und Notwendigkeit eine verschiedene Dynamik vor allem in dem *outer circle* vorweist. Demzufolge ist das irische Englisch als die älteste ‘language-shift‘ Varietät in Europa auch näher dem Britischen Englisch einzuordnen und somit gehört diese Varietät zum *inner circle*. Einen ganz anderen Status weist das jamaikanische Englisch auf, da es durch das Creole-Kontinuum beeinflusst wird,

wohingegen das indische Englisch als eine klassische ESL (Englisch als Zweitsprache) betrachtet wird. Diese werden zum *outer circle* eingeordnet.

Die Analyse zeigt nicht nur quantitative Präferenzen im Gebrauch dieser Konstruktionen sondern vor allem subtile Unterschiede im Gebrauch jeweils in formelle und informelle Gesprächssituationen, ins Besondere im jamaikanischen und indischen Englisch. Wie erwartet, erscheint das Modalverb *must* am häufigsten im jamaikanischen und indischen Englisch im Vergleich zu den anderen drei Varietäten, was üblicherweise als Weiterleben einer konservativen Sprachnorm in diesen Varietäten interpretiert wird. Gleichzeitig ist *have to* die am häufigsten verwendete Konstruktion in allen Varietäten, aber vor allem im *outer circle*, wobei *(have) got to* hauptsächlich in dem *inner circle* zu treffen ist. Interessanterweise ist *need to* am häufigsten im jamaikanischen Englisch und am wenigsten im indischen Englisch verbreitet. Das irische Englisch ist dem britischen Englisch am ähnlichsten. Es lässt sich nur vermuten, dass die hohe Anzahl von *have to* im *outer circle* als eine fortlaufende und im Prinzip alle Varietäten des Englischen umfassende Entwicklung zu erklären ist. Die geringe Häufigkeit der informellen Konstruktion *(have) got to* im *outer circle* könnte auf präskriptivem Einfluss während der frühen kolonialen Phase hindeuten oder als Beweis dafür gesehen werden, dass in der späteren Phase der Kolonisierung britische Neuerungen nicht mehr bereitwillig aufgenommen wurden. Es ergibt sich, dass das amerikanische Englisch die am meist entwickelte Varietät ist, gefolgt vom britischen Englisch, wobei das indische Englisch am konservativsten erscheint.

Ein unerwartetes Ergebnis ist, dass stilistische Merkmale, z.B. Kontraktionen (‘*ve to, ‘ve got to, gotta*), die auf einer Änderung hinsichtlich dem informellen Gebrauch aufweisen würden, im *outer circle* kaum erscheinen und nur sporadisch im irischen (‘*ve to*), britischen (‘*ve got to*) und amerikanischen Englisch (*gotta*) zu finden sind. Dieses lässt sich wahrscheinlich auf die Transkriptionsmethoden der Korpora zurückführen. Ein weiterer Vergleich von ca. 20 geschriebenen Zeitungsartikeln zeigt, dass die hohe Frequenz von *must* im *outer circle*, vor allem im jamaikanischen Englisch, als ein Zeichen für die geschriebene Norm des britischen Englisch einzuordnen ist.

Was den semantischen Kontrast betrifft, gibt es in allen Varietäten einen starken Unterschied zwischen dem epistemischen und deontischen Gebrauch dieser vier Konstruktionen. Dabei ergibt sich ein weiteres unerwartetes Ergebnis, dem zufolge sowohl deontisches als auch epistemisches *must* am häufigsten im ICE-India zu finden sind.

Die wichtigste Beobachtung ist, dass innerhalb der deontischen Notwendigkeit der Unterschied zwischen subjektivem und objektivem Gebrauch deutlich auf eine fortlaufende Änderung in Richtung von *subjectification* (Goossens 2000; Traugott 1989; Traugott & Dasher 2002) in den New Englishes, vor allem im jamaikanischen und indischen English, hinweist. Zudem stellte sich heraus, dass *must* in diesen zwei Varietäten kaum prototypisch als Ausdruck von *speaker-imposed obligation* gebraucht wird, sondern öfters als Ausdruck von *self-imposed obligation* vorkommt. Es bleibt jedoch offen, in wie weit dieser Prozess fortgeschritten ist, da sich die beiden Gebrauchsweisen oft schwer klassifizieren lassen.

Insgesamt zeigt die Analyse, dass drei jüngere Kontaktvarietäten des Englischen eine unterschiedliche Sprachdynamik im Vergleich zum britischen Englischen vorweisen.

Die Arbeit ist in 8 Kapiteln eingeteilt, inklusive Einleitung und Schlussfolgerungen. Kapitel 2 liefert einige Vorbemerkungen zur Terminologie (z.B. New Englishes vs. World Englishes) und zu den einschlägigen Tendenzen in der linguistischen Forschung im Bezug auf die verschiedenen Varietäten des Englischen. Anschließend, wird in Kapitel 3 kurz die historische sowie die sprachliche Entwicklung des standard jamaikanischen, indischen und irischen Englisch erläutert. Kapitel 4 beschäftigt sich mit dem theoretischen und methodologischen Hintergrund zur Analyse des Gebrauchs von Modalkonstruktionen in den drei obenerwähnten Varietäten. Eine detaillierte Darstellung der relevanten Literatur zum Thema soll die wichtigsten diachronen und synchronen Tendenzen in den einzelnen Standard- und nicht-Standardvarietäten des heutigen Englisch beleuchten. Kapitel 5 zeigt signifikante Unterschiede in der quantitativen Distribution aller morphosyntaktischen Merkmale der vier Modalkonstruktionen. Kapitel 6 erweitert den Fokus der Untersuchung auf den semantischen Kontrast dieser Konstruktionen mit dem Ziel den deontischen Gebrauch näher zu beschreiben. Schließlich werden die Modalverben auch auf ihren epistemischen Gebrauch beschrieben. Kapitel 7 befasst sich mit einer statistischen multivariaten Analyse der unterschiedlichen Faktoren. Mehrere morphosyntaktische, semantische sowie pragmatische Variablen werden miteinander untersucht. Hauptziel einer solchen detaillierten Mikroanalyse ist, die Tendenzen der sprachlichen Entwicklung anhand interner Variation dieser Konstruktionen zu beleuchten. Die Ergebnisse der Korpus-Analyse werden in Kapitel 8 zusammengefasst und im Kontext der bisherigen Forschung zu den New Englisches einbezogen. Dazu werden einige weitere Forschungsrichtungen vorgeschlagen.