PAULA RAUTIONAHO

Variation in the Progressive

A Corpus-based Study into World Englishes

■

ACADEMIC DISSERTATION
To be presented, with the permission of
the Board of the School of Language, Translation and Literary Studies
of the University of Tampere,
for public discussion in the Auditorium Pinni B 1096,
Kanslerinrinne 1, Tampere, on November 29th, 2014, at 12 o'clock.

UNIVERSITY OF TAMPERE

PAULA RAUTIONAHO

Variation in the Progressive

A Corpus-based Study into World Englishes

Acta Universitatis Tamperensis 1997
Tampere University Press
Tampere 2014

UNIVERSITY
OF TAMPERE

ACADEMIC DISSERTATION
University of Tampere
School of Language, Translation and Literary Studies
Finland

The originality of this thesis has been checked using the Turnitin OriginalityCheck service in accordance with the quality management system of the University of Tampere.

Cover design by
Mikko Reinikka

Distributor:
kirjamyynti@juvenes.fi
http://granum.uta.fi

Acta Universitatis Tamperensis 1997
ISBN 978-951-44-9635-6 (print)
ISSN-L 1455-1616
ISSN 1455-1616

Acta Electronica Universitatis Tamperensis 1485
ISBN 978-951-44-9636-3 (pdf)
ISSN 1456-954X
http://tampub.uta.fi

Suomen Yliopistopaino Oy – Juvenes Print
Tampere 2014

Acknowledgments

As is evident, there are people to be thanked, people without whom these words would never have been written.

My most sincere gratitude is to Professor Juhani Klemola, my supervisor, my academic guide, my friend. Your unwavering trust in me and my abilities has taken me far, far further than you probably realise. Your encouraging words of criticism have guided me through these years, and the way you have made me feel special is beyond what I ever wished for. Thank you.

I thank my preliminary examiners, Professor Terttu Nevalainen and Professor Marianne Hundt, for their invaluable and insightful advice on how to shape this piece of research further. All remaining errors are mine.

Langnet, the Finnish Graduate School of Language Studies, is to be most sincerely thanked not only for the funding of this thesis, but for offering a wide variety of courses, seminars, and colleagues, some of whom have become friends. Being able to share thoughts on our research, and on matters far removed from research, has been refreshing and rewarding. Izabela Czerniak and Minna Korhonen, thank you for your friendship.

Everyone at the School of Language, Translation and Literary Studies at the University of Tampere deserve to be thanked for creating an environment where it is possible and, even more importantly, enjoyable, to do research. In particular, I would like to thank Professor Päivi Pahta for her never-ending enthusiasm and energy to get things done, and Head of Administrative Affairs Tiina Harjula for knowing everything there is to know about working at the University of Tampere. Fellow doctoral students Hanna Parviainen and Paul Rickman share with me an interest in World Englishes, corpora, and doing the best we possibly can. Thanks for being there, guys! Dr Elina Ranta has walked the same paths as I have, both the path of academia and the path of motherhood: sorry for disturbing you with too long coffee breaks (without the coffee) and thank you for being a friend.

I am lucky to have been a part of two successful research projects, first the GlobE consortium and then the ChangE consortium, both led by professors Juhani Klemola, Markku Filppula and Anna Mauranen. The co-operation of colleagues from three different universities has been supporting and encouraging.

Special thanks go to Joensuu: Professor Markku Filppula shared with me his extensive knowledge on Irish English, while doctors Heli Paulasto and Lea Meriläinen share my enthusiasm in words ending in *ing*.

Over the past years, I have been fortunate to have met (in person or via email) a number of scholars, who have been willing to share their time and knowledge with me. I have received invaluable help from Dr Erik Smitterberg and Professor Douglas Biber. I would also like to thank anyone else who sacrificed their time and effort to ask questions, give advice, and comment on my work: you are too numerous to mention by name.

Having thanked my academic family, it is now time to thank my everyday family. Without them I would have been lost in the somewhat distorted world of doing research, and forgotten about the real world. My parents-in-law, Sisko and Allan Rautionaho, have for more than ten years now been an inseparable part of my family, offering countless hours of care and love for the little ones (and myself). My mother, Hilkka Suoniemi, has given me all I ever asked for. Thank you for the love and care, and the patience. My brothers, Markku and Niko, also deserve to be thanked, just for being there. I only wish that my father, whom we lost far too early, would have seen who I have become.

Thanks are due to Sani the dog, for the wet kisses and the long walks.

There is one man who kept me sane throughout this process, who reminded me that food and sleep are essential to any aspect of life. Harri, your love, patience, and faith in me, your at times doubting words of encouragement kept me going. You are an exceptional man, so kind, so wise, so determined to take care of your family. Thank you for being the best husband I could ever have and for being the best father our daughters could ever have, thank you for being there for me every day. I love you.

Finally, there are two little girls who mean everything to me. Saana and Malla, I know you will grow up to be beautiful, wise, courageous women, who will be able to achieve anything you wish. In the meantime, I want you to enjoy being tiny, without a care in the world. Saana, I promise you, one day you *will* be a grown-up and you can eat anything you like (just remember to brush your teeth). Malla, your smile and laughter make my heart ache; you simply are the happiest girl I have ever met. You two are the love of my life, the light of my day, the essence of my being. I love you both so very much.

This work is dedicated to my family: without you I am nothing.

Tiivistelmä

Englannin kielen progressiivi, eli kestomuoto, muodostetaan BE-apuverbin ja partisiipin preesens -muodon (*-ing*) yhdistelmällä, esimerkiksi *I'm reading Treasure Island*. Progressiivi ilmaisee yleensä parhaillaan käynnissä olevia tapahtumia, mutta sillä voidaan ilmaista myös monia muita asioita, mm. tulevaisuuteen sijoittuvia tapahtumia (*When's your Mum coming back?)*, epävarmuutta (*I'm hoping to go at Easter*) tai staattista tilannetta (*he was standing very close behind us)*. Progressiivin käyttö englannin kielessä on lisääntynyt huomattavasti 1600-luvulta jatkuen nykypäivään, ja lisääntyvän käytön myötä sen muodot ja merkitykset ovat muuttuneet. Esimerkiksi Kranich (2010) tarjoaa laajan kuvauksen muutoksista progressiivin käytössä. Viime vuosikymmeninä progressiivia on tutkittu laajalti, mutta mielenkiinto on kohdistunut lähinnä brittienglantiin, tai britti- ja amerikanenglannin vertailuun.

Englannin kieli on levinnyt Brittein saarilta ympäri maailman: englantia puhutaan paitsi äidinkielenä (esim. Yhdysvalloissa), myös toisena kielenä (esim. Singaporessa) tai vieraana kielenä (esim. Suomessa). Englannin aseman määrittely monikielisessä yhteisössä voi olla haasteellista; virallisesta asemasta huolimatta englannilla saattaa olla vain rajoittunutta todellista käyttöä, tai toisaalta englanti voi samankin yhteisön sisällä olla osalle puhujista äidinkieli ja toisille toinen tai jopa vieras kieli. Näistä ristiriidoista huolimatta englannin levinneisyyttä kuvataan usein nk. kehämallilla, jossa *sisäkehä* viittaa perinteisiin englanninkielisiin maihin; *ulkokehä* maihin, joissa englannilla on virallinen asema toisen tai toisten kielten rinnalla; ja *laajeneva kehä* maihin, joissa englannin asema kansainvälisen kommunikaation kielenä on huomioitu mm. koulutusjärjestelmässä.

Tässä tutkimuksessa tarkasteltiin progressiivin käyttöä maailmanenglanneissa, kolmessa sisäkehään ja viidessä ulkokehään kuuluvassa maassa. Aikaisempien tutkimusten perusteella progressiivin käytön odotettiin eroavan sisäkehään ja ulkokehään kuuluvissa varieteeteissa, mm. esiintymistiheyden ja erityismerkitysten osalta. Tällä tutkimuksella haluttiin paikata progressiivin tutkimuksesta uupuvaa yhtenäistä maailmanenglantien tarkastelua: yksittäisten ja toisistaan erillisten tutkimustulosten vertaaminen toisiinsa ei välttämättä tuota todellista kuvaa progressiivin käytöstä maailmanenglanneissa.

Tutkimuksessa tarkasteltiin neljää osa-aluetta: progressiivin esiintymistiheyttä, morfosyntaktista vaihtelua (mm. aikamuotoja, modaalisten apuverbien esiintymistä), leksikaalista vaihtelua (mm. progressiivissa esiintyvät verbit ja niiden semanttinen luokittelu), ja progressiivin funktioita (mm. prototyyppinen progressiivi, subjektiivinen progressiivi). Koska progressiivi voi esiintyä vain verbiasemassa, on mielekkäämpää verrata progressiivin määrää verbilausekkeiden määrään kuin sanamäärään; tämä vertailu mahdollisti mielenkiintoisten vertailujen tekemisen mm. aikamuotojen (preesens/imperfekti) esiintymisen suhteen. Tutkimuksen aineistona käytettiin *International Corpus of English* -korpuskokoelmaa (seitsemän varieteettia) sekä *Santa Barbara Corpus of Spoken American English* -korpusta. Korpuksista rajattiin tutkimusaineistoksi 100 000 sanaa vapaata keskustelua, sillä progressiivin on osoitettu esiintyvän yleisimmin juuri puhutussa kielessä. Aineiston yhteensä 800 000:sta sanasta haettiin WordSmith Tools 6.0 Concordancer -työkalun avulla BE-apuverbin ja *ing*-päätteen yhdistelmät, ja näistä rajattiin manuaalisesti jäljelle varsinaiset progressiivi-muodot, yhteensä 6 345 kappaletta.

Tämän tutkimuksen perusteella progressiivin käyttö ei määrity varieteetin aseman mukaan, toisin sanoen ulkokehään kuuluvissa varieteeteissa ei voida havaita yhdenmukaista käyttöä joka eroaisi sisäkehän varieteeteista. Paremminkin, progressiivin käytössä on havaittavissa eroja varieteettien välillä, riippumatta siitä mikä englannin asema siinä on. Toisaalta, varieteettien väliset erot ovat suurelta osin melko pieniä. Progressiivin esiintymistiheys on lähes sama kuudessa kahdeksasta varieteetista (noin 610 progressiivia 10 000 verbilauseketta kohden); hongkonginenglannissa (ulkokehä) vastaava luku on selvästi alhaisempi (360), ja irlanninenglannissa (sisäkehä) korkeampi (790). Progressiivin esiintyminen statiiviverbien yhteydessä (esim. *I'm standing*) on usein yhdistetty ulkokehän varieteetteihin: tutkimukseni osoittaa, että statiivisia progressiiveja esiintyy kaikissa kahdeksassa varieteetissa, mutta niiden epästandardi käyttö (viittaaminen pysyviin tilanteisiin, esim. *you must be having waterfalls?*) on rajoittunut lähinnä intian- ja hongkonginenglanteihin. Statiiviverbien epästandardi käyttö näyttäisi olevan kytköksissä myös habituaalisten progressiivien epästandardiin käyttöön: näitäkin esiintyy lähinnä intian- ja hongkonginenglanneissa. Nämä esimerkit osoittavat, että aikaisempien tutkimusten esittämä jako sisä- ja ulkokehiin progressiivin käytön suhteen ei ole täysin paikkansapitävä: progressiivin esiintymistiheys ei ole korkeampi ulkokehän varieteeteissa ja toisaalta ulkokehään yhdistetyt epästandardit käytöt eivät koske kaikkia ulkokehän varieteetteja, vaan rajoittuvat kahteen varieteettiin.

Morfosyntaktisen vaihtelun tarkastelun mukaan progressiivi esiintyy yleisimmin aktiivin preesensissä (70% kaikista progressiiveista), ilman modaalista apuverbiä (95%) tai ajan adverbiaalia (82%), päälauseessa (67%), ja BE-apuverbi on useimmiten kokonainen (54%). Progressiivi on jossain määrin leksikaalisesti rajoittunut: yleisimmin progressiivissa esiintyvät pääverbit ovat hyvin pitkälti samat eri varieteeteissa (GO, DO ja SAY), ja verbien esiintyminen progressiivissa on sidoksissa verbin semanttisiin ominaispiirteisiin. Tutkimuksen mukaan dynaamiset verbit esiintyvät usein progressiivissa, kun taas harvoin progressiivissa esiintyvät verbit ovat enimmäkseen statiivisia. Progressiivin funktioista prototyyppinen käyttö on yleisin jokaisessa varieteetissa (ka 39%), joskin varieteettien väliset erot voivat olla suuriakin (BrE 52%, IrE 29%). Suurin erityisfunktioiden luokka on subjektiiviset progressiivit (18%), joilla voidaan ilmaista mm. omaa suhtautumista tapahtumiin (esim. paheksuntaa), kohteliaisuutta, tai joita voidaan käyttää kerronnallisina tehokeinoina. Tulevaisuuteen viittaavat progressiivit ovat erityisasemassa singaporenenglannissa, jossa niiden osuus on lähes 20% kaikista progressiiveista: tämä mielenkiintoinen seikka kaipaa lisätutkimusta, kuten myös epästandardien statiivisten ja habituaalisten progressiivien käyttö intian- ja hongkonginenglanneissa.

Yksittäisistä varieeteista irlanninenglanti erottuu monessa suhteessa muista: progressiivin käyttö on hyvin yleistä ja erityisesti funktiot jakaantuvat eri tavoin kuin muissa varieteeteissa. Irlanninenglanti on saanut vaikutteita iiristä ja muista kelttiläisistä kielistä, jonka seurauksena sen progressiivi on muokkautunut omanlaisekseen. Samoin mm. intian- ja singaporenenglanneissa voidaan nähdä vaikutteita paikallisista kielistä (mm. hindi; kantoninkiina). Muita mahdollisia syitä progressiivin käytön eroavaisuuksiin ovat puhekielistyminen, kieliopillistuminen sekä kielenoppimiseen liittyvät seikat.

Kaiken kaikkiaan tämä tutkimus osoittaa, että progressiivin käyttö maailmanenglanneissa vaihtelee varieteettikohtaisesti, eikä varieteetteja voida ryhmitellä sen perusteella kuuluvatko ne sisä- vai ulkokehään. Tutkittujen varieteettien ja kieliopillisten piirteiden määrä tekee tästä yhden laajimmista progressiivia käsittelevistä tutkimuksista, ja tutkimuksen metodologiset ratkaisut (progressiivin vertailu ei-progressiiviin) tekevät siitä erityisen kiinnostavan laajemmassakin mittakaavassa.

Avainsanat: progressiivi, aspekti; englannin kieli, maailmanenglanti; korpustutkimus

Contents

Figures

Tables

1 Introduction

This is a corpus-based, descriptive study of the progressive (BE + V*ing*) in World Englishes: the study seeks to address the long-standing interest in the frequency and the use of the progressive and add to the discussion the worldwide dimension. To the present day, only a handful of articles have discussed the progressive in more than one or two (usually British and/or American) varieties of English. This study thus seeks to provide a large-scale comparison of a number of features related to the progressive in a number of different varieties of English spoken throughout the world.

Neither the progressive nor World Englishes are new as objects of scholarly interest. The progressive has attracted interest for close to hundred years, if Jespersen's 1931 account is considered one of the first widely-acknowledged scholarly works on the use of the progressive. Throughout the decades, the progressive has continued to intrigue researchers due to the fact that its use has increased tremendously during the last four centuries, resulting in changes in the form and functions of the progressive. While the progressive is acquiring new forms (e.g. *had been being Ved*, see e.g. Mair 2006), it is, perhaps more importantly, also acquiring new functions (e.g. interpretative progressives, see e.g. Kranich 2010). The extension of the progressive into stative and habitual situations is a widely recognized feature, usually related to Outer Circle, i.e. non-native, varieties of English (see e.g. Mesthrie 2008).

World Englishes, on the other hand, are a more recent phenomenon as an object of study within the field of linguistic variation. The early 1980s witnessed a paradigm shift in the study of English; the focus gradually widened from BrE and AmE to the varieties of English spoken in other parts of the world. New journals were launched to serve scholars interested in the identification and description of global varieties of English: *English World-Wide* (1980-), *World Englishes* (1981-), and *English Today* (1985-). Studies on World Englishes aim to describe individual varieties of English and to investigate how a variety differs from others as regards, for instance, the lexis, syntax, or pragmatics. On the other hand, studies on World Englishes may focus on a feature of grammar, lexis, or phonology and seek to establish how this particular feature is used in different varieties of English. The

present study is of the latter kind: the aim here is to investigate the frequency and functions (among other features) of the progressive in altogether eight World Englishes.

A few terminological remarks are needed to set the scene for this study. First, the term 'the progressive' is used to refer to the combination of any form of the verb BE and the present participle of a verb, i.e. V*ing*; the main object of interest in this study is *'m reading* in *I've read even The Hobbit and I'm reading Treasure Island*.[1] The progressive has been called by many names, e.g. 'the continuous (aspect)', 'temporary aspect', or 'the expanded form', but as 'the progressive' is widely used and acknowledged it is used in the present study, as well.

As regards the term 'World Englishes', however, the matter of terminology is slightly more complex. We need to specify the use of 'World' rather than 'New' or 'Postcolonial', and the use of the plural 'Englishes' rather than 'English' in the singular. In addition to 'World Englishes', there are a number of other terms used to describe the new orientation towards global varieties of English; these include 'New Englishes' and 'Postcolonial Englishes', among others.[2] The term 'New Englishes' was first used by Platt, Weber and Ho in the early 1980s, and refers to localized forms of English spoken in parts of Asia, West and East Africa, and parts of the Caribbean. According to Platt, Weber and Ho (1984: 2-3), New Englishes

a) have developed through the education system, as a subject to be taught and as a medium of instruction,

b) have developed in an area where most of the population spoke a language other than a native variety of English,

c) are used for a number of functions, e.g. in parliament, government, media, and literature; and

d) have adopted features of their own and thus become localized.

While some scholars found the term useful, others criticized the fact that it referred to something that was 'new' mainly to Western scholars (Schneider 2007: 3).

The early 1980s saw the advent of another term, i.e. 'World Englishes', introduced by Braj Kachru and Larry Smith as they took over the editorship of a journal they renamed *World Englishes* in 1985 (Bolton 2003: 3). The new term

[1] Source: ICE-GB, S1A-013.

[2] McArthur (2004) provides a summary of the terminology related to the study of English as a world language.

avoids the caveat of drawing a dividing line between varieties regarded 'old' or 'new'. The term 'Postcolonial Englishes', on the other hand, was introduced in the early 21st century by Edgar Schneider (2003, 2007): Schneider wanted to offer a more neutral perspective on varieties of English, devoid of political opinions. Schneider sees varieties of English as "products of a specific evolutionary process tied directly to their colonial and postcolonial history" (2007: 3).[3]

The plural form 'Englishes', rather than the singular 'English', adds stress to the fact that English no longer has one major variety that all speakers of the language consider the standard, a source of authority and prestige (Mesthrie and Bhatt 2008: 3). While the term 'World English', in the singular, is used to refer to "the idealised norm of an internationally propagated and internationally intelligible variety of the language" (Bolton 2003: 4), 'World English*es*', on the other hand, refers to "localised varieties of English used intranationally in many 'ESL' societies throughout the world" (ibid.).

In the present study, the term 'World Englishes' is understood to involve all varieties of English throughout the world, including also BrE and AmE, which are sometimes set apart from other varieties as the reference points. Thus, this study follows the thinking of McKay (2010), among others, in that the World Englishes paradigm is regarded as placing "all varieties of English on par with each other without any one being a reference point" (McKay 2010: 91). In other words, BrE and AmE, the two major varieties of English, are considered equal to any other variety of English – they just happen to have a long and influential standing among varieties of English spoken throughout the world. The differences between World Englishes arise from the different paths of evolution they have undergone (see Schneider 2007). Throughout this study, comparisons are drawn between the individual varieties and between Inner and Outer Circle Englishes – a method widely used in previous studies on World Englishes.

The present study is one of the first large-scale comparisons of the use of the progressive in World Englishes. Answers are sought to the following research questions:

a) What is the frequency of the progressive in World Englishes: e.g. is the progressive more frequently used in Outer Circle varieties?

b) What is the grammatical environment of the progressive: e.g. are there differences in the verb phrase patterns in different varieties of English?

[3] Schneider's Dynamic Model is further discussed in Chapter 3 below.

c) Which main verbs is the progressive used with: e.g. are the verbs shared in World Englishes?

d) What are the functions of the progressive in World Englishes: e.g. do Outer Circle Englishes extend to non-standard uses?

A number of research questions of smaller scale can be attached to each of the main questions, as has been exemplified above.

The outline of the present study is the following: Chapter 1 has introduced the background and the aims, after which Chapters 2 and 3 discuss the theoretical background of the study. Chapter 2 focuses on the progressive: matters related to aspect are briefly discussed in more general terms, and previous studies into the progressive are surveyed, after which the functions in which the progressive is used in present-day English are presented. Chapter 3 discusses the concept of World Englishes and provides an introduction into each of the eight varieties included in the study. The methodological background of this study is presented in Chapter 4, in which the aims of the study are discussed in more detail, and the data and the retrieval process are carefully described. The core of the present study, the use of the progressive in World Englishes, is divided into four components, each discussed separately in Chapters 5 to 8. Chapter 5 begins the unraveling of the progressive in World Englishes with a discussion on the frequency of the construction: the three sections of Chapter 5 present the frequency of the progressive in previous studies as well as in the present data, and discusses the results in detail. Chapter 6 focuses on various morphosyntactic features related to the progressive; these are tense, perfect progressives, modal auxiliaries, and the progressive passive, as well as clause type, contracted forms and temporal adverbials co-occurring with the progressive. Chapter 7 discusses lexical choices related to the main verbs of progressive verb phrases, while the functions of the progressive in World Englishes are presented and discussed in Chapter 8. Finally, Chapter 9 summarises and discusses the results of the study.

2 The progressive

2.1 The progressive as an aspectual marker

The term 'aspect' is used to refer to the "different ways of viewing the internal temporal constituency of a situation" (Comrie 1976: 3f.). The *perfective* aspect makes a claim about the situation as a whole, including the endpoints, while the *imperfective* aspect focuses on a middle part of the situation, excluding the endpoints. As aspect in English has been discussed in numerous publications, only a short discussion of the most important aspectual features related to the progressive in English is included here. For a more detailed account of aspect, see e.g. Comrie (1976), Brinton (1988), Smith (1991), Binnick (1991), or Ziegeler (2006). Also Kranich (2010) discusses aspect and aspectual choices at length, and provides a visualization of the two aspectual options, which she has adapted from Klein (1994) (see Figure 1 below). The plus signs indicate a situation taking place, the short lines indicate that the situation no longer holds, and the brackets indicate the time for which an utterance makes an assertion. As Figure 1 portrays, the perfective views the situation as a whole (as indicated by the square brackets), while the imperfective concentrates on some middle part of the situation excluding the beginning and the ending of the situation.

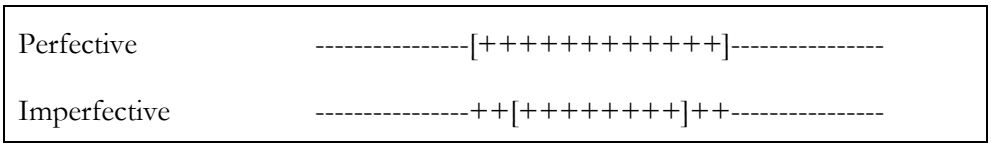

Figure 1. Visualization of aspectual meaning (Kranich 2010: 28).

According to Smith (1991: 5), sentences contain two kinds of aspectual information: *situation type*[4] is signalled by the verb and its arguments, while *viewpoint*

[4] A term related to the situation type is 'Aktionsart' or 'lexical aspect', which refers to "the intrinsic temporal qualities of a situation" (Brinton 1988: 3). Vendler's (1957) Aktionsart classification is

is expressed by grammatical morphemes attached to the verb. Additional temporal information is provided by tense and adverbials (ibid.). Smith (1991: 12) points out that it is the speaker who makes the choices on which viewpoint and situation type to use; aspect is thus subjective. The perfective and the imperfective, discussed above, are the main types of viewpoint. Situation types, on the other hand, are made up of different combinations of the properties of dynamism, durativity and telicity (Smith 1991: 6). States (e.g. *know the answer*) are static and durative, while Activities (e.g. *laugh*) are durative and atelic, i.e. they do not have a natural end-point. Accomplishments (e.g. *build a house*) and Achievements (e.g. *reach the top*) both have an end-point which is the outcome of a change of state. The two situation types differ in that Achievements are punctual, while Accomplishments are durative. Finally, Semelfactives (e.g. *cough*) are punctual atelic events. Situation type is initially determined by the verb, but the presence of prepositions, adverbials, or other forms may lead to changes in the situation type portrayed by the verb phrase. For instance, an Accomplishment situation such as *write a report* becomes an Activity if a durative adverbial is added: *Jerry wrote a report for two hours* (Smith 1991: 71, 158). Thus when determining the situation type of a sentence, one should take into account not only the verb but other elements, as well. Figure 2 (based on Smith (1991: 30)) below summarizes the five situation types.[5]

Situation type	Dynamic	Durative	Telic
State	-	+	-
Activity	+	+	-
Accomplishment	+	+	+
Semelfactive	+	-	-
Achievement	+	-	+

Figure 2. Features of the situation types.

According to Smith (1991: 222), the main subtype of the imperfective is the progressive, which is typically durative and dynamic and thus available to non-

probably the most well known, on which later systems are often based. For more information on Aktionsart, see e.g. Binnick (1991) or Smitterberg (2005).

[5] Sometimes more than these three parameters are included to enable more accurate analysis of the situations. Vendler's (1957) original classification included the parameter [+/-] Voluntary, while Brinton (1988) adds [+/-] Homogeneity and [+/-] Multiplicity. For the purposes of the present study, the three parameters draw an accurate enough picture.

stative events. Activities combine with the progressive most easily and retain the semantic properties of the progressive (e.g. *Joan is singing*). When Accomplishments occur with the progressive, the end-point is typically excluded from the situation and the focus is on the activity in progress prior to reaching the goal (e.g. *Joan is singing a song*) (van Rooy 2006: 42). Quirk et al. (1985: 208) point out that the goal is not necessarily reached when Accomplishments occur in the progressive. Activities and Accomplishments are characterized as having duration and are thus compatible with the progressive. Semelfactives and Achievements, on the other hand, do not have duration but are punctual; they thus require some re-analysis when used in the progressive. Atelic punctual events, i.e. those without an end-point, combine with the progressive to express a series of repeated events (e.g. *John was nodding his head*) (ibid.). With telic punctual event, on the other hand, the progressive focuses the attention on the period of time leading up to the change of state (e.g. *The train is arriving at platform 4*) (p. 209). As regards States, some of them can occur in the progressive but are then reinterpreted as temporary or as carrying "emotional color" (e.g. *John is believing in ghosts*) (Smith 1991: 225-6). The situation types in connection with the progressive are further discussed in section 2.3 below.

The meaning of the progressive remains a source of debate. As Binnick (1991: 281) points out, in spite of the fact that speakers of English agree that the progressive differs from the simple forms, no one has succeeded in completely or satisfactorily specifying what exactly is the meaning of the progressive. The difference between the progressive and the corresponding simple form may be subtle, as in *it's looking like rain/it looks like rain*, or more radical, as in *he is eating lunch in the café/he eats lunch in the café* (ibid.). Although numerous authors have discussed the meaning or meanings of the construction at great length, consensus seems to have been reached mainly on the fact that the progressive has something to do with duration, which is evident in the names used of the construction. The progressive aspect has also been called *durative* aspect, *continuous* aspect, or *temporary* aspect (e.g. Quirk et al. 1985: 197). Another term used is *expanded form* (e.g. Jespersen 1931). The term used in this study is *the progressive* (occasionally *the progressive form*), as it seems to be the most widely used term in recent research.

A situation referred to by the progressive is typically viewed as incomplete and in progress (Bybee et al. 1994: 125-6; Comrie 1976: 24). Quirk et al. (1985: 197-8) exemplify the differences between the simple and the progressive aspect with the following sentences:

(2.1) Joan *sings* well.

(2.2) Joan *is singing* well.

In (2.1), the simple present indicates Joan's ability to sing in general, while (2.2) refers to a particular occasion or season of Joan's singing. The past tense forms of the same sentences (consider (2.3) and (2.4) below) have slightly different meanings. The simple past in (2.3) views the situation as a whole, while the past progressive in (2.4) indicates that the activity was in progress and incomplete.

(2.3) Joan *sang* well.

(2.4) Joan *was singing* well.

Leech (1987) uses three dimensions with which to separate the progressive from the non-progressive: *duration* (stretches the time-span of an event)*, limited duration* (compresses the time-span of a state) and *not necessarily complete*. Leech exemplifies the three features with the following sentences (pp. 19-20):

Duration*:*

(2.5a) The house *falls* down!	(sudden movement)
(2.5b) The house *is falling* down.	(gradual movement)

Limited duration*:*

(2.6a) Which team *do* you *support*?	(generally speaking)
(2.6b) Which team *are* you *supporting*?	(at this particular match)

Not necessarily complete*:*

(2.7a) The man *drowned*.	(the man died)
(2.7b) The man *was drowning*.	(the man did not necessarily die)

The difference in sentences (2.5a) and (2.5b) is the fact that the progressive stretches out the time the house takes to fall down, or in other words, slows down the event time. The effect of the progressive is the opposite in sentences (2.6a) and (2.6b): the present simple refers to a more or less permanent habit, while the progressive, in this case, compresses the habit to a temporary one, concerning a particular moment. Finally, the event portrayed by the progressive need not be complete; it is possible that the man in sentence (2.7b) was saved. Leech's third feature is related to telic events, i.e. Accomplishments and Achievements.

Kranich (2010: 30-37, 169-172) discusses the possibility that the progressive in present-day English could be evolving into a general imperfective marker. This possibility arises from the fact that the progressive can be used to describe situations other than dynamic and ongoing, as well, i.e. stative and habitual

situations. Kranich refers to a number of previous studies which provide more or less successful answers to the question, and based on her own data comes to the conclusion that, although there is a rise in the proportion of progressives used as markers of general imperfective aspect, the rise is more likely connected to the overall rise of progressives and cannot be regarded as evidence that the progressive is developing into a general imperfective marker (p. 172).

2.2 Survey of previous studies

The progressive has been extensively studied in the past decades. The increasing frequency of the form has attracted a number of scholars to investigate the diachronical changes of the progressive (e.g. Smitterberg (2005), Kranich (2010)), while particular grammatical features related to the progressive have also acted as the source of interest (e.g. Hundt (2004a) and Smith & Rayson (2007) on the progressive passive, or Depraetere & Reed (2000) on the present perfect progressive). Table 1 below lists studies on the progressive: the topics of these studies indicate that the progressive has indeed been widely researched but mainly in BrE. In the following, three of the studies listed in Table 1 are introduced in more detail. These three studies are Römer (2005) on spoken present-day BrE, Smitterberg (2005) on written and spoken 19th-century BrE, and Kranich (2010) on the progressive in Modern BrE. Studies on the progressive in World Englishes (Collins 2008, Hundt and Vogel 2011, and van Rooy 2014) are discussed separately in section 3.4 below.

Table 1. Previous studies on the progressive.

Author(s) and publication year	Title	Focuses on
Collins (2008)	The Progressive Aspect in World Englishes: A Corpus-based Study.	Inner and Outer Circle varieties: various aspects
Hilbert & Krug (2012)	Progressives in Maltese English. A comparison with spoken and written text types of British and American English.	MalteseE, BrE, AmE; various aspects
Hundt (2004a)	Animacy, agentivity, and the spread of the progressive in Modern English.	BrE: spread of the progressive
Hundt and Vogel (2011)	Overuse of the progressive in ESL and learner Englishes - fact or fiction?	12 varieties representing IC, OC and EC: stative progressives
Kranich (2010)	The Progressive in Modern English. A Corpus-based Study of Grammaticalization and Related Changes.	BrE: various aspects
Leech et al. (2009)	The progressive.	BrE, AmE: various aspects
Mair and Hundt (1995)	Why is the Progressive Becoming More Frequent in English. A Corpus-Based Investigation of Language Change in Progress.	BrE, AmE: spread of the progressive
Meriläinen, Paulasto & Rautionaho (fc.)	Extended uses of the progressive form in Inner, Outer and Expanding Circle Englishes. (working title)	11 varieties representing IC, OC and EC: stative and habitual progressives

Nesselhauf (2007)	The spread of the progressive and its 'future' use.	BrE: futurate use
Römer (2005)	Progressives, Patterns, Pedagogy: A Corpus-Driven Approach to English Progressive Forms, Functions, Contexts and Didactics.	BrE: various aspects
van Rooy (2006)	The extension of the progressive aspect in Black South African English.	BSAfE: meaning and use
van Rooy (2014)	Progressive aspect and stative verbs in Outer Circle varieties.	IndE, KenE, BSAfE: stative uses
Schilk & Hammel (2014)	The progressive in South Asian and Southeast Asian varieties of English – mapping areal homogeneity and heterogeneity.	BrE and seven OC varieties: lexical aspects
Smith (2002)	Ever Moving on?: The Progressive in Recent British English.	BrE, AmE: various aspects
Smith (2005)	A Corpus-based Investigation of Recent Change in the Use of the Progressive in British English.	BrE: various aspects
Smitterberg (2005)	The Progressive in 19th-century English: A Process of Integration.	BrE: various aspects
Williams (2002)	Non-Progressive and Progressive Aspect in English.	BrE: progressive vs. non-progressive

Römer (2005). The first half of Römer's book is a very detailed account of the contexts and uses of the progressive in two present-day BrE corpora, while the second half addresses how progressives are taught in German schools. Römer adopts an approach to the data that is slightly different from most other studies on the progressive, i.e. *corpus-driven* linguistics. Corpus-driven linguistics starts from the data and postulates theories based on observations elicited from the data, rather than applying pre-existing theories on corpus data. Römer includes a detailed account of previous studies on the progressive, but the analysis of her data arises purely from the corpus data without pre-existing hypotheses or expectations. The study discusses the contexts that the progressive occurs in[6], the uses of the progressive and the co-occurrence of the progressive and verbs. At times it feels that the features investigated are not directly related to the progressive itself, but to the verbs that occur in the progressive (e.g. prepositions) and thus slightly irrelevant. Nevertheless, Römer provides valuable information on the use of the progressive in present-day English in a fresh, corpus-driven manner.

The second part of the book compares the real BrE use of the progressive and the class-room reality in foreign language teaching in Germany, and shows that there is a mismatch between the two types of English. Some features of the progressive are over-represented in the teaching material (e.g. use of temporal adverbials with the progressive), while others are under-represented (e.g. of progressives with past time reference) (pp. 271-3). Römer goes on to suggest how teaching materials should be altered to take into account the real use of the progressive represented in corpora. All in all, Römer's investigation into the contexts and uses of the progressive and the discussion on pedagogical implications of the study offer valuable information on the progressive and, on the other hand, on the methodology of corpus research in more general terms.

Smitterberg (2005). In his 2005 book, Smitterberg provides detailed information on the frequency, grammatical environment and functions of the progressive in BrE in the 19th century. The study further includes a valuable methodological discussion: how exactly the progressive should be defined and its frequency measured. The study focuses on a number of linguistic and extralinguistic features related to the progressive, and thus provides a wide-scale picture of how the progressive was used between the years 1800 and 1900, an era which has proved to be of importance to the spread of the progressive. Smitterberg's findings indicate

[6] The features investigated are tense, contraction, subject and object of progressive verb phrases, prepositions, negation, clause type and adverbial specification.

that the 19th century is crucial for the increase of the progressive – the frequency nearly doubled during the century as the feature became more fully integrated into English grammar (p. 245). However, Smitterberg points out that the increase did not take place consistently in English as a whole; rather there were, for instance, genres in which the progressive became clearly more frequent while some genres seemed to be unaffected by the changes in the use of the progressive (p. 244). The results further indicate that changes took place in the formally simple contexts rather than in the more complex environments[7] (p. 246). Regarding lexical variation of the progressive, Smitterberg notes a similar tendency to avoid complexity: the type/token ratio for progressive main verbs decreased (p. 246). An important factor related to the use of the progressive in the 19th century is the fact that the percentage of stative verbs co-occurring with the progressive began to increase (p. 246). Smitterberg argues that the so-called 'not-solely-aspectual' types of progressives[8] may be responsible for the increase in the use of the progressive in general as the frequency of potentially experiential progressives increased significantly during the 19th century (p. 246). Results from Smitterberg's study are discussed in more detail in Chapters 5 to 8.

As regards the present study, Smitterberg's investigation provides valuable points of comparison: as this study was largely inspired by Smitterberg's study (methodological choices, choice of features to be investigated), it is natural that the results from the two studies be compared. The period of time under investigation in Smitterberg's study was a time of colonial expansion, and many of the varieties included in this study have their roots in 19th-century English.[9] It will thus be interesting to see what the results from the two studies reveal about the English spoken, on the one hand, in 19th-century England and, on the other hand, in present-day World Englishes.

Kranich (2010). Kranich (2010) provides a wide and detailed account of the progressive in Modern English, i.e. from the year 1600 to 1999, thus making it the most extensive study on the progressive to date. In addition to the vast time span, the study investigates a large number of both linguistic and extralinguistic features, and includes a thorough theoretical discussion on the functions of the progressive.

---

[7] The progressive passive is an exception: its frequency increased considerably during the 19th century (Smitterberg 2005: 246).

[8] Not-solely-aspectual progressives are more or less comparable to what Kranich (2010) calls subjective progressives, and are further discussed in section 2.3 and Chapter 8 below.

[9] The varieties and their history are discussed in section 3.3 below.

Kranich's findings largely support those obtained by Smitterberg (2005), but are able to provide information on a considerably larger time span, especially as Kranich briefly discusses the use of the progressive prior to 1600, as well. The progressive is favoured in certain genres, is increasingly used in the more simple constructions (except for the progressive passive), and the number of stative verbs co-occurring with the progressive starts to rise during the 19th century. Kranich's study shows that some developments related to the progressive date back centuries, while others are more recent. As regards the increasing frequency of the progressive, Kranich offers a twofold explanation, according to which the rise of the aspectual functions explains the overall increase in the 17th and 18th century. However, the 19th and 20th century developments are better explained by the establishment and increasing use of the interpretative progressive.[10] In the present study, Kranich (2010) is used as an important source of information regarding the functions of the progressive. Although the categorization presented by Kranich is not strictly applied to the present data, it is still used as the backbone of the semantic analysis in this study.

In addition to studies such as the three introduced above, i.e. separate articles published in journals or in collections of articles, and whole books dedicated to the progressive, grammars of English need to address the progressive. While most grammar books include at least a short note about the progressive, the most exhaustive accounts are presented in Quirk et al. (1985: 197-213) and Biber et al. (1999: 470-5). Both Quirk et al. and Biber et al. discuss the use of the progressive at great length, and from various angles including the forms it is used in, the main verbs that it occurs in, and the functions it is used in. Biber et al. (1999), in particular, acts as a solid reference point for the present study, as it includes quantitative results in addition to the description of the use of the progressive. Biber et al. (1999) have also influenced the contents of the study, as Chapter 7 on the lexical variation of the progressive is largely based on matters discussed therein (pp. 471-5).

[10] The interpretative progressive is introduced in section 2.3 and further discussed in section 8.4.

2.3 Functions of the progressive in present-day English

This section discusses the functions in which the progressive is used in present-day English. How exactly these functions transpire in World Englishes is discussed in Chapter 8. The framework for the functions of the progressive used in the present study is largely based on Kranich (2010). Table 2 below provides a categorization of progressive functions included in this study. There are five main functions – progressive, habitual, stative, subjective, and futurate – and some of these are further divided into subgroups.

Table 2. Functions of the progressive in present-day English.

Function	Subgroups	Code
Progressive		PROG
Stative	1) Delimited stative (standard)	PSTA
	2) Non-delimited stative (extended)	ESTA
Habitual	1) Delimited habitual (standard)	PHAB
	2) Non-delimited habitual (extended)	EHAB
Subjective	e.g. attitude, emotion, tentative, interpretative	SUB
Futurate	1) *will*/*shall* + *be* V*ing*	FUT1
	2) Futurity evident from context	FUT2
Other	(e.g. WORK-verbs)	OTH
Indeterminate		IND

A basic division is made between aspectual uses (progressive, habitual, stative, futurate[11]) and subjective uses, as these are identified as the two main types of progressive functions by Kranich (2010). According to Kranich (2010: 168), an aspectual progressive focuses on "some middle-part of the situation not including its endpoints", while a subjective progressive conveys "speaker attitude, emphasis or the speaker's subjective interpretation". In some cases, the only difference between the progressive and the non-progressive is subjective rather than aspectual (ibid.: 62); when an aspectual reading of a progressive sentence seems improbable, a subjective reading may well be in order. The main functions of the progressive are discussed below.

The core sense of the progressive is typically defined as temporariness, i.e. when a situation is referred to with the progressive it is considered to be in existence for a limited period of time (see e.g. Quirk et al. 1985: 197-8; Biber et al. 1999: 470). Another key concept related to the progressive is the fact that the situation is considered to be in progress and also incomplete (ibid.). In the present study, a prototypical progressive is considered to portray events as being in progress and incomplete at the time of utterance (following e.g. Quirk et al. 1985: 199). Thus, sentences such as (2.8) and (2.9) below are considered to portray the prototypical or core use of the progressive. In (2.8), the situation, i.e. watching television, is in progress at the time of utterance (UT), has begun before UT and will go on after UT. Similarly, in (2.9), the engine was in operation at UT, and the situation took place before and after UT.

(2.8) Mary *is watching* television. (taken from Quirk et al. 1985: 199)

(2.9) The engine *was running* smoothly. (taken from Quirk et al. 1985: 207)

As pointed out in section 2.1, when the progressive is combined with different situation types, the focus may shift. Activities, such as *watching* and *running* above, occur with the progressive without problems. When the progressive is combined with Accomplishments, the focus is on the activity in progress prior to reaching the goal, as in (2.10) below. Similarly in (2.11) below, the attention is directed at the period of time leading up to the change of state (i.e. when the weather has warmed

[11] Collins (2008: 240) regards futurate uses of the progressive as non-aspectual and thus examples "of the grammaticalization of the progressive aspect". Leech et al. (2009: 133), on the other hand, consider futurate progressives to be extensions of the basic meaning of the progressive, i.e. aspectual. In the present study, futurate progressives are considered aspectual rather than non-aspectual.

up). In both examples, the end-point is excluded from the situation, and it is practically irrelevant whether the goal is reached or not.

(2.10) Jill *is knitting* herself a sweater. (taken from Quirk et al. 1985: 208)

(2.11) The weather *is getting* warmer. (taken from Quirk et al. 1985: 207)

Punctual situation types, as opposed to Activities and Accomplishments, do not have duration and are thus incompatible with the progressive without re-analysis. When Achievements co-occur with the progressive, the resulting interpretation is similar to that of Accomplishments, i.e. the focus lies on the period leading up to the change of state (consider (2.12) below). With Semelfactives, on the other hand, the re-interpretation is that of repetition: a typically punctual act, such as *fire*, is seen as a series of repeated acts when used in the progressive (consider (2.13) below).

(2.12) The queen *was dying.* (taken from Quirk et al. 1985: 209)

(2.13) Someone *was firing* at us. (taken from Quirk et al. 1985: 209)

While the four situation types described above do occur rather freely with the progressive, albeit some of them requiring re-analysis, States are a more complicated matter. Generally, the progressive does not combine with verbs that are used statively, as States do not change, i.e. there is no element of progression (Quirk et al. 1985: 198, see also Declerck 1991: 167). The progressive expresses a dynamic reading of a situation, whereas a state consists of segments similar to one another. When stative verbs do occur in the progressive, they imply that the situation is *temporary* rather than permanent (Quirk et al. 1985: 198). Thus, the situation in sentence (2.14a) below is interpreted as permanent, whereas sentence (2.14b) refers to temporary residence:

(2.14a) We *live* in the country.

(2.14b) We *are living* in the country.

Table 3. Semantic classes of stative verbs.

Category	Verbs
Being/having	*be, have, cost, require,* etc.
Cognition/emotion/attitude	*think, feel, forget, remember,* etc.
Perception/sensation	*see, hear, smell, hurt, taste,* etc.
Stance	*sit, stand, live, lie,* etc.

Stative verbs may be categorized into four different groups (following Huddleston and Pullum 2002 and Leech 1987; see Table 3 above). Verbs of being and having (as in (2.15) below) refer to permanent situations which do not have a dynamic reading, and are thus incompatible with the progressive (Declerck 1991: 169). Verbs of cognition, emotion or attitude (as in (2.16) below) are not normally used in the progressive, but may do so when they express temporariness or tentativeness (Quirk et al. 1985: 203). Verbs of perception and sensation (as in (2.17) below) form a more complex group in which some verbs are compatible with the progressive while others are not. Verbs referring to involuntary reaction of the senses (e.g. *see, hear*) are not usually used in the progressive (Declerck 1991: 167), while with verbs of bodily sensation (e.g. *hurt, itch*) the non-progressive and the progressive are practically interchangeable (ibid.: 169). Finally, stance verbs (as in (2.14) above) are 'intermediate between the stative and dynamic categories' and can be used in the non-progressive, to express a permanent state, and in the progressive, to express a temporary state (Quirk et al. 1985: 205). A number of verbs belonging to one of the stative categories may be used with the progressive – such exceptional uses are discussed by Declerck (1991: 169-174) and Quirk et al. (1985: 203-5).

(2.15)	He *has* a nice wife.	(**is having*)[12]
(2.16)	I *love* you.	(?*am loving*)
(2.17a)	I *noticed* her leave the shop.	(**was noticing*)
(2.17b)	He said his stomach *hurt*.	(*was hurting*)

Stativity should be thought of as a cline rather than a dichotomy, as some stative verbs yield to being used in the progressive more easily than others. Qualities (often referred to by verbs of being and having: *be tall, have two legs*) are the most permanent type and incompatible with the progressive, while other verbs belonging to the same category (*be angry/ill, love, resemble*) are, according Quirk et al. (1985: 200), less permanent and may, under specific circumstances, be used in the progressive; they then adopt a dynamic reinterpretation. For instance, in sentence (2.18) below, 'awkwardness' is considered deliberate behavior rather than a

[12] *Have* may be used as activity verb, as in *He's having a good time* (Declerck 1991: 168). *Have* is further discussed in section 7.1 below. Examples (2.15) to (2.17) are taken from Declerck (1991: 167-9).

permanent trait. Similarly, in (2.19), *being tired* must be interpreted as pretense rather than an actual state.

(2.18) Peter *is being* awkward.

(2.19) Mary *is being* tired.

At the other end of the scale we find stative verbs that are incompatible with the progressive. When a speaker despite the semantic restrictions decides to use such a verb in the progressive, the use is regarded as *extended*. Extended use of stative progressives refers to instances in which the situation described is not temporally delimited, i.e. temporary, but "convey a sense of extended duration" (van Rooy 2014: 157). Van Rooy (2014) discusses three types of stative progressives, and ascertains that non-standard use of stative progressives is a purposeful extension of the progressive paradigm (see section 3.4 below).

When the progressive combines with the *habitual*, the resulting meaning is that of repetition taking place over a limited period of time (Quirk et al. 1985: 199). Sentences (2.20a) and (2.20b) below (taken from Quirk et al. 1985: 199) exemplify the difference between the simple and the progressive habitual; in the simple form, the habit is thought of as permanent, while the progressive (and the temporal clause) render the habit temporary:

(2.20a) The professor *types* his own letters.

(2.20b) The professor *is typing* his own letters while his secretary is ill.

Leech (1987: 32) points out that the presence of a temporal adverbial or temporal clause is not necessary; it is the temporariness of the habit that allows the use of the progressive. For instance, the period of time implied in (2.21b) below is shorter than that implied in (2.21a):

(2.21a) I *take* dancing lessons.

(2.21b) I*'m taking* dancing lessons.

Similarly to stative uses of the progressive, there are two kinds of uses of the habitual progressive: those that follow the semantic restrictions laid out by the progressive in Standard English (henceforth StE), and those that violate the restrictions and are thus non-standard. Standard use of the habitual progressive is

characterized by the use of a temporal adverbial or temporal clause to indicate that the habit takes place over a limited period time[13], whereas non-standard use of the habitual progressive results in habitual situations that are not limited in time. Habitual progressives and their extended use are discussed in section 8.5 below.

As pointed out above, when an aspectual reading of a progressive verb phrase fails, a *subjective* reading may be more suitable. Consider sentences (2.22a) and (2.22b) below (taken from Kranich 2010: 62):

(2.22a) Paul *always sleeps* at our apartment.

(2.22b) Paul'*s always sleeping* at out apartment.

Both (2.22a) and (2.22b) describe a habit, but the progressive does not render the situation portrayed in (2.22b) dynamic or ongoing. In fact, the difference between sentences (2.22a) and (2.22b) lies in the fact that the speaker is expressing his or her attitude towards the situation: the speaker is annoyed by Paul sleeping at the apartment. This type of subjective progressive is called, by Kranich (2010: 63), *subjective progressive with ALWAYS*[14], and is discussed, for instance, by Leech (1987), Huddleston and Pullum (2002), Smitterberg (2005) and Leech et al. (2009). The consensus seems to be that, normally, subjective progressives with ALWAYS imply *negative* speaker attitude, but, as Kranich (2007) points out, the functions may also carry positive (see (2.23) below) or neutral (see (2.24) below) attitude.

(2.23) I'*m always enjoying* your work because you*'re constantly bringing* something new to the plate. Keep at it.[15]

(2.24) The universe *is forever expanding*.

Sentences such as (2.24) represent progressives that are not exactly aspectual (nor subjective), as the event or act referred to goes on unchanging for an unlimited period of time (*forever*).

As this type of the subjective progressive includes an overt marker, i.e. an adverbial of time, it is relatively easy to spot such instances of the progressive.

---

[13] As mentioned above, it is also possible to indicate limited duration without explicit temporal specification; limited duration is then evident from context.

[14] ALWAYS refers to all adverbials with the meaning 'always', i.e. *always, constantly, forever*, etc.

[15] Sentence (2.23) comes from Kranich's (2007) web-based study which shows that subjective progressive with ALWAYS clearly more frequently conveys negative attitude (52 out of 92 instances) than positive attitude (16 out of 92 instances) (Kranich 2007: 131, Table 1).

However, the progressive may carry subjective function without the presence of ALWAYS. Smitterberg (2005: 221) suggests formal criteria to distinguish subjective uses from the aspectual, but Kranich (2010: 206-13) points out a number of problems related to these criteria.[16] The process of locating subjective progressives in the present data is based on careful reading of the data: as there are no clearcut criteria to distinguish subjective progressives from aspectual progressives it is evident that the analysis is far from objective.[17]

A number of subjective progressives without ALWAYS are identified by Kranich (2010: 62, 66-68): the progressive may be used to indicate the speaker's attitude, tentativeness, vividness or the progressive may be used as a narrative device. Uses of subjective progressives without ALWAYS are illustrated in the following examples, taken from Kranich (2010: 62, 67-68, 221):

(2.25) If you have a good estate, every covetous rogue *is longing* for it.

(2.26) I'*m* (just) *wondering* whether he'll come.

(2.27) Then they*'re stuffing* all morning and you slave away, you*'re rushing* to get this Christmas dinner.

(2.28) …so she starts singing in Norwegian and I *am* just *cracking* up thinking this is some joke that someone's played and you know people *are* just *looking around* like what is this […] so they stopped…

In (2.25), similarly to subjective progressives with ALWAYS, the speaker uses the progressive to express his or her attitude towards the situation described. In (2.26), the use of the progressive renders the utterance more tentative than if the non-progressive (*I wonder whether…*) was used, and the tentativeness may be further strengthened by using the adverbial *just*. In (2.27), the progressive is used to describe "the situation in a more vivid, more graphic way" in order to get the addressee's attention (Kranich 2010: 67). Finally, sentence (2.28) is an example of a narrative use of the progressive[18]: the progressive here does not have aspectual

[16] Accoding to Smitterberg's (2005: 221) criteria 'potentially experiential' (i.e. subjective) progressives should include at least three of the four criteria: present tense without perfect or modal auxiliary, main clause, first- or second-person subject, and a stative situation. Kranich (2010: 206-13) discusses the matter at lenght and comes to the conclusion that the formal criteria suggested by Smitterberg (and others) are not sufficiently reliable to locate all and only subjective progressives.

[17] Comparing the distribution of subjective progressives obtained by different scholars is problematic, as the criteria used may differ to a great extent.

[18] Such use of the progressive is called the 'foregrounded progressive' by Couper-Kuhlen (1995).

meaning, but should rather be regarded as a narrative device drawing attention to specific events described. Such use of the progressive is found in conversational AmE, but not in written BrE (Kranich 2010: 68).

Yet another type of subjective progressive is the interpretative progressive, used to "interpret – or give meaning to – a situation with which the addressee is assumed to be familiar, either because it is mentioned explicitly (as in (2.29) below), or inferable from the context (as in (2.30) below)" (Leech et al. 2009: 134).

(2.29) When Paul Gascoigne says he will not be happy until he stops playing football, he *is talking* rot.

(2.30) You'*re kidding*!

In (2.29), the *when*-clause expresses the situation in question, while the progressive is used to provide the speaker's interpretation of what such a situation means to him or her personally. Following Ljung's (1980: 70f.) description of interpretative progressives, the *when*-clause constitutes part A, in which the simple form is used to express the observed behavior, while the progressive constitutes part B, interpreting the behavior expressed by part A. Kranich's findings (2010: 222) indicate that the increasing frequency of the interpretative progressive may account for the general increase in the frequency of the progressive; similar findings are reported by Leech et al. (2009) and Mair (2006).

Futurity, in combination with the progressive, may be indicated by three means; first, by the combination of *will* or *shall* and the progressive[19] (see (2.31) below), second, by the progressive combined with a temporal adverbial or subclause that indicates a future time reference (see (2.32) below), and, third, futurity may be implied in the context without a temporal adverbial (see (2.33) below).

(2.31) "We *shall be suffering* a critical handicap in the election if the trade union can't help us on this," he urged. (Kranich 2010: 183)

(2.32) I *am leaving* for New Zealand next Thursday. (Nesselhauf 2007: 195)

(2.33) We'*re not taking* you home! (Nesselhauf & Römer 2007: 313)

Leech (1987) describes the use of *will* or *shall* in combination with the progressive as expressing a time-frame in the future or as referring to an event happening

[19] Modal auxiliaries *will* and *shall* may occur in combination with the progressive indicating epistemic or deontic function rather than futurity. The discussion of modal auxiliaries in section 6.3 includes all instances of *will* and *shall*.

"independently of the will or intention of anyone concerned", i.e. 'future as a matter of course' (Leech 1987: 68). Sentence (2.31) is, according to Kranich (2010: 183), an example of 'natural course of events' dependent on the condition in the *if*-clause, rather than on *our* will or intention. Kranich's findings indicate a rise in the frequency of *will/shall* + *be* X*ing* in the late 20th century, although the overall frequency remains low (2010: 186). The second and third means of referring to the future with the progressive are grouped together under subgroup 'futurity evident from context', as there are no structural markers such as a modal auxiliary present. The presence of a temporal adverbial or a temporal subclause from which futurity is evident is a clear sign and easily recognizable. However, in order to recognize a progressive referring to the future when there is no co-occurrence of temporal adverbials or subclauses one needs to consider another ways of signaling futurity. According to Nesselhauf and Römer (2007: 313), such signals include conditional clauses or noun phrases indicating futurity, and negation (see (2.33) above).[20] At times, it is also possible that temporal adverbials are in 'removed context', i.e. occur at a distance from the progressive; in such cases, the future reference remains clear to the participants of a conversation (in a spoken interaction), even though the reference is not explicitly repeated (ibid.). Nesselhauf and Römer (2007) report that, in informal conversation in particular, progressives with future reference occur more often *without* than *with* temporal specification.

Functions considered to be marginal, i.e. few in number but clearly functioning differently from the five main categories of progressives, are included in category 'Other'. Another group of progressives categorized separately are the so-called WORK-verbs. According to Brinton (1988: 199), verbs such as WORK and STUDY indicate either a series of events or a permanent state and are thus semantically ambiguous. Further, Paulasto (2014: 259) points out that these verbs are so frequent in ICE-IND, for instance, that they might distort the results of the study if included.[21]

How exactly the functions of the progressive described above are attested in present-day World Englishes is discussed in Chapter 8. Before that, the progressive is thoroughly investigated from a number of points of view: Chapter 5 discusses the frequency of the construction in World Englishes, Chapter 6 focuses on the morphosyntax of the progressive, and the topic of Chapter 7 is lexical variation of

[20] When it is clear that a negated progressive does not refer to the present, it may refer to the future (Nesselhauf & Römer 2007: 313).

[21] Instances included in the 'Other' category are nevertheless discussed in Chapter 8 below.

the progressive. At times, it may be impossible to separate issues that would, as a matter of fact, belong to another category of investigation dealt with in another chapter. This is the case with stative verbs, for instance, which are discussed in Chapters 7 and 8. Stative verbs are, strictly speaking, a lexical phenomenon as it is the verb that is inherently active or stative. On the other hand, stative verbs form one of the functional categories that the progressive is used in, and therefore are discussed in Chapter 8, as well. This evidently results in some overlap in these two chapters, but excluding discussion on stative verbs from either chapter does not seem like a viable option.

Before the results of the present study are discussed, the varieties included in the study need to be identified - this is done in the following chapter. After that, Chapter 4 discusses the methodological choices made, introduces the data and the retrieval process of progressives.

3 World Englishes

3.1 English as a world language

The story of English as a world-wide language begins in the late 16th century, when the first colonies outside the British Isles were established. One could argue, as Leith (2007: 118) does, that the colonial times began even earlier than the 16th century, within the British Isles themselves, as the English language supplanted the Celtic languages spoken in Ireland, Scotland and Wales. According to this line of thinking, also IrE and AmE are considered post-colonial varieties, albeit with a considerably longer history and a more stable status than varieties spoken in Southeast Asia, for instance.[22]

The reasons behind the English desire to explore and colonize land far away from the British Isles are numerous, one of the most important ones being the desire to grow into a world power. Other European states, e.g. Portugal and Spain, already had a head start in acquiring possessions outside the state borders (Leith 2007: 118), and so the English also set off in search of new land and subjects. According to Leith (ibid.: 119-120), the motivations for establishing colonies were economic, social or political; in addition to gaining access to great wealth, the colonies presented a way of escaping problems at home. Unemployment, population growth, or religious animosity could be left behind and a new life begun overseas.

Although the processes of colonization differ from one case to another, it is possible to discern three types of colonization: *displacement*, *subjection* and *replacement* (Leith 2007: 120). In the first case, the precolonial population was *displaced* by the English-speaking settlers (e.g. America), while in the second case, the precolonial population was *subjected* to the settlers' control and a small group of the native population was granted access to English to enable them to work as interpreters, etc. (e.g. India). The third type of colonization involves the *replacement* of the precolonial population by new labour transported elsewhere (e.g. Jamaica). The

[22] Schneider (2007: 251-308) provides a case study of AmE as a post-colonial variety.

linguistic outcomes of the three types of colonization differ from one another; displacement of the original population led to gradual disappearance of the native language(s) and English becoming the first language of the new nation, while in the case of subjection, the native language(s) continued to be spoken and English was introduced as a second or additional language. In countries where the precolonial population was replaced by people transported from a number of origins, pidgins and creoles emerged (Mesthrie and Bhatt 2008: 20).

The British Empire initiated and supported the spread of English throughout the world, but as the Empire started to crumble with more and more colonies gaining their independence, one of the former colonies took over and secured the position of English as an international lingua franca and a world language: the United States of America (Sedlatschek 2009: 1). Since the Second World War, the United States has grown into a global superpower on the economic and political fronts, and perhaps even more importantly, on the media and mass communication front. American music and film industries, in addition to the internet, have secured English as the language of the world (Tottie 2002: 245). Indeed, Mair (2013) suggests that the global importance of AmE has superceded that of BrE: Mair's theoretical model 'World System of Standard and Non-Standard Englishes'[23] characterizes AmE as a 'hyper-central variety', the "hub" of World Englishes, while BrE is labelled as a 'super-central variety' (p. 264).

The total number of people who use English is difficult to estimate, not least due to difficulties in obtaining the data: there is no single source of information providing data on the number of speakers of English. Crystal (2003) resorts to a number of different sources and to estimations based on a percentage of the population in instances where no official information is available (ibid.: 61-65). Crystal points out a number of other problems related to this matter, one of the greatest being the fact that English is used and viewed differently in each country (ibid.: 66-67). English may be an official or a co-official language of a country, or it may hold no official status although it is used extensively. Crystal points out that varieties of English are usually thought of as existing on a continuum from a standard form of English to a creole form of English, and that these varieties are not always mutually intelligible (ibid.). Depending on what exactly is considered

[23] Mair (2013) introduces a model for handling "uses of English in domains beyond the post-colonial nation state" (p. 253), according to which both standard and non-standard varieties may be characterized as hyper- or supercentral, central or peripheral varieties. For instance, IrE is a standard central variety, and BrE a standard super-central variety.

English, the number of speakers may differ greatly. Despite these difficulties Crystal does provide a list of countries and the number of speakers of English as L1 or L2, and comes to the total of approximately 760 million speakers of English around the world (ibid.: 62-5).[24] Crystal points out that this figure should be considered a minimum: the true total of L1 and L2 speakers is bound to be larger as not all countries where English is spoken were included in the list, and as the true percentage of L2 speakers may be greater than what Crystal used as the estimate (pp. 67-8). In addition to L1 and L2 speakers, the majority of people who use English, either daily or less often, have learned the language at school as an additional language. Crystal (2003: 61) estimates that there are anywhere between 500 million and 1 billion such users of English, a fact which makes English a truly global language.

When talking about English as a world language, we must take into account the fact that English is not uniform, but varies from one location to another. Trudgill (2004) proposes three stages through which new dialects evolve.[25] During stage 1, settlers accommodate their individual speech patterns in order to be able to interact with other settlers with different speech patterns by eliminating features that are not regular or shared by the majority of the community (ibid.: 89). A new dialect starts to take form during stage 2, as the children of the first settlers acquire a new way of speaking by choosing features from a pool of competing features originated from the dialects of the previous generation (p. 101).[26] A new dialect, distinct from the ones used by the previous generations, stabilizes during stage 3 (p. 113). An important concept related to new-dialect formation is *drift*, i.e. if dialects derive from a common source, they are likely to evolve along similar lines (see Trudgill 2004: 129-133). However, as Trudgill and Hannah (2002: 6) point out, linguistic change takes place in all languages and dialects regardless of where they originate: varieties of English keep evolving due to mechanisms of linguistic change at different levels of language. Another important factor affecting varieties of English the world over is language contact: as speakers of English come into contact with speakers of indigenous languages, the languages are bound to have an effect on each other, again on different levels including the lexis as well as grammar (ibid.: 8).

---

[24] According to Crystal, there are approximately 330 million people who use English as L1 and 430 million as L2 (2003: 65).

[25] Trudgill's study is based on the emergence of NZE, but it is reasonable to assume that the stages are similar regarding other varieties, as well.

[26] The feature pool theory is also discussed by Mufwene in various papers (e.g. Mufwene 2001).

Thus varieties of English (as well as of any other language) are constantly undergoing change due to internal as well as external forces.

The following section discusses theoretical models with which World Englishes are classified into more easily approachable entities.

3.2 Conceptualizing World Englishes

Previous research on varieties of English mainly relies on two models with which to conceptualize World Englishes, based on the functional and political role of English in a given country. The first of these models, suggested by Strang (1970) and Quirk et al. (1972), distinguishes between 'English as a Native Language' (ENL), 'English as a Second Language' (ESL), and 'English as a Foreign Language' (EFL). Schneider (2007) provides a brief summary of the status of English in the three categories. According to Schneider (2007: 12), in ENL countries, English is the first language of all or the majority of the population (e.g. Great Britain, Australia), while in ESL countries, English is used widely and often in official functions with one or more indigenous languages used simultaneously (e.g. India, Nigeria). In EFL countries, English is learned through formal education and, while it has no official standing, it is used widely in some domains, for instance, the press (e.g. Egypt, Taiwan) (ibid.). Schneider (ibid.) points out that while the model is widely used, it fails to take into account the complexity of linguistic situations in countries such as South Africa, or of non-native-speaking groups in ENL countries (e.g. Native Americans), or of native speakers of English in ESL countries (e.g. in Singapore).

The second model widely used in research on varieties of English, and building upon the ENL-ESL-EFL distinction, is Kachru's "Three Circles" model (Kachru 1985). Kachru argues that his model takes into account "the types of spread, the patterns of acquisition and the functional domains in which English is used" (ibid.: 12), thus providing more information than the earlier model. The three concentric circles that varieties are classified into convey information on the status of a variety of English in a given country, as well as on the political history of a variety of English (ibid.). According to Kachru (ibid.), the *Inner Circle* includes countries in which English is the primary language (e.g. Great Britain, the United States of America, Australia and Canada), while in countries outside the Inner Circle, English is not the primary language but a secondary or an additional language. Countries that were once colonized by the British and where the English language

remains as a legacy of the colonial times, are part of the *Outer Circle* (e.g. India, Nigeria, Singapore). Outer Circle countries are characterized by multilingualism: in India, for instance, English is an associate official language, and in Singapore, English is one of four official languages (ibid.). According to Kachru, English has penetrated a number of levels of the Outer Circle societies. English has a number of domains in which it is used, both inter- and intranationally, and it is used by speakers at different levels of the society, not only at the higher levels (ibid.). The third and outermost circle, called the *Expanding Circle*, includes countries in which English does not necessarily have an official status or a colonial history, but which nonetheless recognise the importance of English as an international language (such countries include, for instance, Japan and Russia) (ibid.). While the earlier model based on the ENL-ESL-EFL distinction largely emphasized the importance of ENL, Kachru's model places more importance on the Outer and Expanding Circles (Schneider 2007: 14). According to Schneider (ibid.), Kachru seeks to emphasize the position of English as a world language, as a language that belongs to anyone who uses it, not only to those who use it as a native language.

Although widespread, both models described above have faced criticism, mainly due to the fact that there are countries that do not easily fit into the models (e.g. Jamaica or Malaysia). In his 2003 article, Schneider proposes a new model for the categorization of varieties of English.[27] This model builds upon the idea that

> speakers keep redefining and expressing their linguistic and social identities, constantly aligning themselves with other individuals and thereby accommodating their speech behavior to those they wish to associate and be associated with (Schneider 2007: 21).

Schneider's model considers the evolution of English varieties to proceed through five phases during which changes in identity building affect linguistic choices made by the participants involved in the "us vs. them" situation, i.e. settler and indigenous groups (p. 29). During Phase 1, named *Foundation*, English begins to be used in a new territory, and Phase 2, *Exonormative stabilization*, follows as English stabilizes as the language of administration, legal system, etc. (p. 33). It is during Phase 2 that the form of English brought to the territory starts to change as a consequence of contact with the local language(s) (pp. 39-40), and a new distinct variety of English is born during Phase 3, *Nativization* (p. 44). The territory and its

[27] The Dynamic Model is further discussed in Schneider's 2007 book *Postcolonial English: Varieties Around the World.*

people undergo a number of changes during Phase 3: political independence is gained, and regular contact between the two groups leads to closer and closer ties between them (p. 41). As English is being used by more and more people with different language backgrounds, the English used begins to change: conservative speakers of English may feel that the language is deteriorating and voice their concerns in the pages of local newspapers (p. 43). These concerns are forgotten during Phase 4, *Endonormative stabilization*, as the new variety with its distinct forms is accepted and used as a means of expressing the new independent identity of the people now unified (p. 49). The fifth and final phase, *Differentiation*, leads to the birth of regional and social dialects of the local variety of English (p. 54).[28] Schneider's Dynamic Model has been noted for its applicability to most postcolonial varieties of English (Mesthrie and Bhatt 2008: 35-6).

A number of other models for handling World Englishes have been suggested by, for instance, McArthur or Görlach, both in the late 1980s. For more information on other models of English, see e.g. Jenkins (2003: 19-21) and Mesthrie and Bhatt (2008: 27-31).

3.3 Varieties included in the present study

In this section, the varieties of English included in this study are briefly introduced. Primarily, matters of interest are the history of English and its current status in each of the countries. The eight varieties are also discussed in the light of the three models discussed above.

British English. Present-day British English[29] is the result of centuries of contact between peoples and languages, including, for instance, the Scots, the Irish and the Welsh; the French; and more recently, people from all over the Commonwealth countries. Due to villages and towns being relatively isolated from one another, and also because people seldom travelled to distant locations, distinct regional varieties emerged (see e.g. Gramley and Pätzold 2004; Wolfram and Schilling-Estes 2006). These traditional dialects differ to a great extent in grammar, morphology, lexis and pronunciation (Gramley and Pätzold 2004: 229).

---

[28] The phases of the Dynamic Model are summarised in tabular form in Schneider (2007: 56).

[29] In the present study, British English is considered to consist of the varieties of English spoken in England, Wales and Scotland, following the make-up of the ICE-GB corpus (see Nelson et al. 2002: 4).

More recently, the traditional, especially rural, dialects have been reported to be declining, whereas the urban dialects are evolving due to urbanization (see e.g. Cheshire 1991: 16). According to Gramley and Pätzold (2004: 229), syntactic differences between major urban centres are levelling out. They further point out that grammatical variation within urban varieties of English is not as much regional as it is educational, so that the more highly educated people are likely to use StE and the less educated use forms closer to the non-standard (ibid.).

A number of languages other than English are spoken in the British Isles. The ones with a historical background in the region are Welsh, Scots, Scottish Gaelic, and Irish, and more recent immigrant languages include Polish, Tamil, Urdu, and Punjabi, for instance.[30] The presence of these languages (as well as other mechanisms of language change) has shaped present-day BrE into a variety different from the one spoken at the time of colonial expansion. Thus in order to truly see the connections between the traditional form of BrE landing into new territories and the new varieties of English spoken the world over, a data set covering language use of the 19th century would be needed. However, as the present study focuses on variation in present-day English, it is more suitable to use recent data. A glimpse into the historical variety of BrE is given in Smitterberg (2005), which is used as a reference point throughout the study at hand.

Table 4. Facts on Great Britain.[31]

Number of residents	63 million
Significant ethnic groups	white (of which English 83.6%, Scottish 8.6%, Welsh 4.9%, Northern Irish 2.9%) 92.1%, black 2%, Indian 1.8%, Pakistani 1.3%, mixed 1.2%, other 1.6% (2001 census)
Languages	English (recognized regional languages: Scots, Scottish Gaelic, Welsh, Irish, Cornish)

Irish English. Although English was introduced to Ireland for the first time in the 12th century, it was in the 17th century that the English language succeeded in

[30] Source: http://www.ethnologue.com/country/GB, accessed on January 9th, 2014.

[31] Source: https://www.cia.gov/library/publications/the-world-factbook/geos/uk.html, accessed on September 19th, 2013.

planting its roots in Irish soil. The Anglo-Normans of the first wave spoke mainly French and the southwestern dialect of English (Barry 1982: 84), and as they were gradually absorbed into the local population, English was to be found only in a small enclave on the east coast (Harris 1991b: 37). In the 16th and particularly 17th centuries, the spread of English into Ireland was more forceful and successful; native speakers of English and Scots took residence in major seaport towns and set up protestant plantations in the eastern and northern parts of the island (Harris 1991a: 195). According to Harris (ibid.), these rural areas in the east and the north were the targets of British colonization in Ireland, and outside these areas contact with the English and their language was considerably weaker. Later on, as English spread to regions outside the plantations, it was transmitted by native speakers of Irish, which resulted in varieties of Irish English that were heavily influenced by the substrate, i.e. Irish (ibid.). Barry (1982: 89) further argues that, as the emerging variety was cut off from mainland Britain, IrE remained archaic and conservative.

English gained more popularity among the Irish population during the 19th century, as famine and emigration more or less halved the population. The implementation of an English-language education system and the adoption of English as the language of the Roman Catholic Church increased the number of English speakers rapidly (Barry 1982: 89; Harris 1991b: 38). By 1900, 85% of the Irish spoke English only, while only 5% of the population claimed to speak only Irish (Barry 1982: 89). Harris (1991b: 38) estimates that in 1970, the proportion of native Irish speakers in Ireland was 2%, and according to Hickey (2008: 74), only a few small enclaves on the western seaboard remain where Irish has an unbroken history as a native language. Recently, language enthusiasts, as well as nationalist-minded Irishmen, have attempted to keep the Irish language alive by using it daily as a second language in otherwise English-speaking urban areas (Hickey 2008: 74). Irish is also being used in broadcasts, government publications and in literature (Leith 2007: 127). Attitudes towards English in Ireland are twofold: on the one hand, English is the native language of the majority of the Irish population, and on the other hand, the Irish language is considered the "carrier of native culture" (Hickey 2007: 22).

Today, the use of English in Ireland is comparable to the situation found in countries such as the United Kingdom and the United States of America. According to Filppula (2008: 328), educated IrE morphology and syntax are similar to those found in other British Isles Englishes; however, more and more differences are found as the form of IrE moves further away from the standard. Filppula (2008: 329) argues that the most important differences between IrE and

other Englishes spoken in the British Isles are found in the tense-aspect-modality system.

Present-day IrE is the language of two politically separate entities: the Republic of Ireland (founded in 1921) and Northern Ireland, which is part of the United Kingdom. In the present study, this division is not taken into account when discussing the use of the progressive in IrE.[32]

Table 5. Facts on Ireland.[33]

Number of residents	4.7 million
Significant ethnic groups	Irish 87.4%, other white 7.5%, Asian 1.3%, black 1.1%, mixed 1.1%, unspecified 1.6% (2006 census)
Languages	English, Irish

American English. The English first set foot on American soil in the late 16th century, and the first colony was established in 1607 in Jamestown, Virginia (Schneider 2007: 254). A number of settlers to other regions on the east coast followed during the 17th century, and gradually the newcomers started to move westward in search of land and both religious and political freedom (ibid.). Svartvik and Leech (2006: 79) divide the settling of the United States into three periods: the Colonial Period covers the years from 1607 to 1790 when all of the colonies had ratified the Constitution; the National Expansion Period (1790-1865) was a time of major population movements, both within the continent (west- and southward) and into the continent (major influx of Irish, Scots, and German immigrants); and finally the Third Period at the end of which, in 1929, changes were made in the immigration laws. Svartvik and Leech (ibid.) point out that the type of immigration changed in the late 19th century. The earlier immigrants were slaves from Africa and free people from the British Isles and Northern Europe, while later on the vast majority of newcomers were from Southern and Eastern Europe. Still later, the North American continent experienced major influx of people of Hispanic and

[32] ICE-IRL is structured so that comparisons between IrE used in the Republic of Ireland and Northern Ireland are possible, cf. Kallen and Kirk (2008) for more information.

[33] Source: https://www.cia.gov/library/publications/the-world-factbook/geos/ei.html, accessed on September 19th, 2013.

Asian origin (Bean and Stevens 2003). Although the volume of immigration was clearly greater during and after the 2nd and 3rd periods, the cultural and political power remained in the hands of the immigrants of the Colonial Period and their descendants (Svartvik and Leech 2006: 79).

The varieties of English spoken in the United States have been influenced by a number of languages spoken by immigrants: French, Spanish, German, and West African languages, as well as Irish and Scots Gaelic (Wolfram and Schilling-Estes 2006: 114). Naturally, also the languages spoken by Native Americans were a source of vocabulary as the settlers came across a multitude of flora and fauna that they had no words for (ibid.). According to Wolfram and Schilling-Estes (2006: 121-2), the westward movement of people led, first, to extreme dialect mixing and, then, to dialect leveling as a result of increasing contact between people speaking different dialects of AmE. In addition to these natural changes, Americans intentionally wanted to distinguish their language from that of the former colonial masters; this was achieved by new coinages and new spellings (ibid.: 115). In the 20th century, the ease of transportation and intercommunication led to increased contact between regional dialects and shifts in what some dialect features stand for: former regional markers have become "markers of social class, ethnicity, and urban-rural distinctions" (ibid.: 128-9). Although regional differences are found in present-day AmE as well, of more importance is variation due to the race and ethnicity of the speaker, as well as to the socio-economic status of the speaker (Gramley and Pätzold 2004: 257). Middle class workers seek to distinguish themselves from the working class with linguistic choices (ibid.), and the use of certain non-standard grammatical features may reveal the speaker's level of education (ibid.: 259). As Gramley and Pätzold point out, the majority of Americans have lost all traces of ethnicity displaying their origin, but a number of ethnic groups may still be distinguished by the way they speak, namely Native Americans, Hispanic Americans[34], and African Americans (ibid.: 259-265).

As pointed out by Tottie (2002: 245), the fact that Americans are dominant in many spheres of life today – be it economics, politics, or entertainment – means that their language and the way they speak it (i.e. AmE) is becoming more and more dominant throughout the world. Although historically speaking, BrE is the 'mother variety' of World Englishes, it may be questioned whether the title should be handed over to AmE.

---

[34] Hispanic Americans may be divided into three major groups: Cubans, Puerto Ricans, and Mexicans (or 'Chicanos') (Gramley and Pätzold 2004: 260-1).

Table 6. Facts on the United States of America.[35]

Number of residents	310 million
Significant ethnic groups	white 63.7%, Hispanic 16.4%, black 12.3%, Asian 4.7%, other 2.9%
Languages	English 80.0%, Spanish 12.2%, Chinese, Tagalog, Vietnamese, French, Korean, German, Russian and others each less 1%

Jamaican English. From the mid-17th century onwards, the British controlled the large sugar plantations on the island of Jamaica, and transported massive numbers of slaves from West Africa to work on the plantations (Schneider 2007: 227). Experience from the earlier Caribbean plantations facilitated the rapid growth of the Jamaican colony, and with the continuous influx of slaves, the number of blacks was soon higher than that of whites (ibid.). The fact that new slaves were usually trained by experienced slaves rather than their white masters led to a situation where the field laborers spoke creole, their masters a form of English, and "an intermediate" class of blacks commanded mesolectal speech forms and were able to communicate with both social groups (ibid.: 229, 232). Slavery lasted until the 19th century, when a number of political changes ended the era and brought education to the former slaves (ibid.: 229). Only after WW2 did Jamaica experience massive changes: democratization, urbanization and socioeconomic diversification, which led to growing sense of nationalism and, finally, independence in 1962 (ibid.: 234).

The official language of Jamaica is English, which in practice refers to "in speech an educated Caribbean accent (…) and in writing an unmarked grammatical common core interspersed with lexical Jamaicanisms" (Schneider 2007: 235). However, the language most commonly used in Jamaica is Jamaican Creole; Devonish and Harry (2008: 277) suggest that most speakers of JamE are, in fact, native speakers of JamC, and that JamE is for most a second language learned in

[35] Source: U.S. Census Bureau, 2010 American Community Survey: http://factfinder2.census.gov/bkmk/table/1.0/en/ACS/10_1YR/DP05, and http://www.mla.org/cgi-shl/docstudio/docs.pl?map_data_results, both accessed on May 6th, 2014.

school. Jamaican Creole, which developed in the late 17th century as blacks outnumbered whites and as newcomers were assimilated into the existing slave population, has traditionally been connected to the more private and relaxed sphere of life, but has recently gained prestige and is used in more and more formal contexts, including politics (Schneider 2007: 233-6). Devonish and Harry (2008: 256) describe the linguistic situation in Jamaica as diglossia: JamC is used for private, informal and oral use, while JamE is for public, formal and written use. According to Devonish and Harry (ibid.), only a minority of Jamaicans are able to use both varieties, with the majority being speakers of JamC.

The diglossic perception of the Jamaican linguistic situation is contested by another theory, i.e. the post-creole continuum. Patrick (2008: 610), citing DeCamp's (1971) original idea, describes the continuum as "a continuous spectrum of speech varieties, ranging from (…) 'broken language' (…) to the educated standard". In the case of Jamaica, the 'broken language', or the basilect, is JamC, and the educated standard, or the acrolect, is JamE. Deuber (2009: 47-8) argues that the practices of speech in Jamaican "can be more appropriately described in terms of a continuum from more English to more Creole ways of speaking", while the written practices can be described as diglossia. However, as Devonish and Harry (2008: 277) point out, even diglossia regards both JamE and JamC as "idealized forms of speech", attesting varying degrees of interaction, rather than as stable and fossilized forms of speech. Whatever the theoretical perception of the Jamaican linguistic situation, the reality is that the language heard at home, at the office, in school, and recently also in the mass media, is JamC rather than JamE. Mair (2009: 59) argues that the influence of the traditional BrE norms on JamE is being replaced by influences from AmE and, more importantly, JamC.

Table 7. Facts on Jamaica.[36]

Number of residents	2.9 million
Significant ethnic groups	black 91.2%, mixed 6.2%, other or unknown 2.6% (2001 census)
Languages	English (official), Jamaican Patois (national)

[36] Source: https://www.cia.gov/library/publications/the-world-factbook/geos/jm.html, accessed on September 19th, 2013.

Indian English. During its four hundred years of existence on the Indian subcontinent, English has grown from a language of only a handful of speakers into a major variety of English. During the 17th century, trading and missionary posts were set up on the Indian coastline by the British East India Company, and English was taught in the missionary schools from early on (Schneider 2007: 162). At first, English was connected to the trade and everyday instrumental use, but by the late 18th century, as the British tightened their grip over India, English became more and more rooted in the country as a language of authority (Bhatt 2008: 547; Schneider 2007: 162). An English-based education system was implemented in the mid-19th century; the majority of the teachers were not native speakers of English, but they were supervised by Scots, Irish and Welsh residents, which may have had an effect on the form of English learned by Indians (Bhatt 2008: 548). The 20th century witnessed a decrease of English-medium teaching, as well as a more general detachment from the British rule (Schneider 2007: 165). India gained independence in 1947, but English remains a firmly established language with co-official status and a multitude of functions.

Today, English has the status of an associate official language, with Hindi as the official language, and 15 national languages being recognized by the Indian Constitution (Sahgal 1991: 299). English is considered a neutral language that poses no threat to any of the numerous ethnic groups of India, whereas the official language, Hindi, being the mother tongue of one ethnic group, continues to meet resistance from native speakers of other local languages (ibid.). English is a language of the government and politics, of higher education, of the law, business and media (Schneider 2007: 167-8). Sahgal's 1991 study showed that, at the time, English was an institutional language, and that educated Indians were beginning to use English more and more often to communicate with their family and friends (pp. 302-3).

Present-day Indian English consists of a number of varieties with a hierarchy from standard to vernacular; while Standard IndE differs from Standard BrE mainly on the phonological level, Vernacular IndE shows syntactic variation as a result of contact with local languages (Bhatt 2008: 546). Schneider (2007: 168) argues for "structural nativization on all levels", and lengthy lists of features considered to be characteristic of IndE are provided in a number of publications (see Balasubramanian 2009: 8). Mukherjee (2007: 179-180) describes a 'triangular model' of IndE; according to this model, IndE consists of a common core of English shared by native and nonnative varieties of English, of interference from local languages, and of autonomous linguistic creativity. According to Mukherjee

(ibid.), the triangular model takes into account variation from acrolectal Standard IndE to basilectal Vernacular IndE, as well as variation from written to spoken IndE. That IndE shows such great level of variation is probably due to the fact that, during centuries, teachers of English have not been native speakers of English, but locals who had learned the language from other locals or native sailors, merchants and missionaries (Schneider 2007: 167).

Table 8. Facts on India.[37]

Number of residents	1.2 billion
Significant ethnic groups	Indo-Aryan 72%, Dravidian 25%, Mongoloid and other 3% (2000)
Languages	Hindi 41%, Bengali 8.1%, Telugu 7.2%, Marathi 7%, Tamil 5.9%, Urdu 5%, Gujarati 4.5%, Kannada 3.7%, Malayalam 3.2%, Oriya 3.2%, Punjabi 2.8%, Assamese 1.3%, Maithili 1.2%, other 5.9%

Philippine English. Of the varieties included in the present study, Philippine English is the only one to have developed as a result of contact with AmE rather than BrE. The United States of America gained control over the area in 1898 as a consequence of the Spanish-American War, and as the new colonial masters regarded their own culture and language superior to the indigenous ones, English was rapidly declared the official language of the nation (Schneider 2007: 140). English became the language of education, of the government, of business and of religion, and was considered a common lingua franca amid the multitude of indigenous languages[38] (Tayao 2008: 292). The spread of English over the region was effective and fast, in part due to the efforts of the American teachers shipped from the US, and in part due to the fact that English was considered "a socioeconomic equalizer", enabling upward movement on the social ladder (Schneider 2007: 140).

[37] Source: https://www.cia.gov/library/publications/the-world-factbook/geos/in.html, accessed on September 19th, 2013.

[38] According to Tayao (2008: 292), there are 87 ethnic languages spoken in the Philippines.

The 20th century witnessed both the rise and fall of English in the Philippines. In the beginning of the century, English was successfully implemented into the Philippine region, and the years of Japanese occupation during WW2 further strengthened the status of English as a sign of resistance (Schneider 2007: 141). After WW2, a bilingual education plan promoting the use of both English and Filipino[39] was implemented in the Philippines, but as Filipino suffered from limited resources, lack of language development and some resistance against its choice as a national language, the status of English remained as strong as ever, if not stronger (ibid.). The end of the 20th century, on the other hand, has marked the end of the golden era of English in the Philippines. According to Schneider (2007: 143), political changes in the country have brought with them a cultural revolution, as a result of which mass culture is now based on local traditions rather than Western. Furthermore, English has become to be seen as the language of the elite rather than the masses, and the knowledge of English has failed to fulfil the expectations associated with it (ibid.). Tayao (2008: 292) reports on "a wave of nationalism": the importance of native languages in education and other formal domains has lead to the decline of English in the Philippines.

Today, English remains a language of formal domains, such as business, politics, higher education and parts of the media, and educated Filipinos still use it extensively (Schneider 2007: 141; Tayao 2008: 293). The national survey of 1993 reports rather high percentages of the population being able to function in English (Thompson 2003: 73)[40]. According to Thompson (2003: 54-5), English is still considered crucial in order to gain access to higher-paid jobs and overseas employment.

Gonzalez (2004: 12) describes PhiE as "monostylistic" and close to the written norms, mainly due to the limited range of domains that it is used in. Furthermore, English in the more informal contexts is being replaced by Taglish, a mixed code of English and Tagalog, i.e. Filipino (Schneider 2007: 142). According to Thompson (2003: 41), Taglish is being used by the majority of educated Filipinos in informal settings to soften the formality of a situation and to express familiarity, not unlike JamC in Jamaica, for instance. Thompson (2003: 41) argues that "Taglish has become Filipino street English", and further discusses "the rapid demise of English" (p. 42) in the Philippines.

---

[39] Filipino is based on Tagalog, one of the native languages (Schneider 2007: 141).

[40] According to the 1993 census, 73% of Filipinos read, 59% write, 74% understand spoken English, 56% speak, and 42% think in English (Thompson 2003: 73).

Table 9. Facts on the Philippines.[41]

Number of residents	105 million
Significant ethnic groups	Tagalog 28.1%, Cebuano 13.1%, Ilocano 9%, Bisaya/Binisaya 7.6%, Hiligaynon Ilonggo 7.5%, Bikol 6%, Waray 3.4%, other 25.3% (2000 census)
Languages	Filipino and English (official)); eight major languages Tagalog, Cebuano, Ilocano, Hiligaynon or Ilonggo, Bicol, Waray, Pampango, and Pangasinan

Singapore English. In 1819, Sir Stamford Raffles established a trading post in what was to become Singapore, and due to the fact that the new port was situated close to already existing trading routes, its economy and population soon started to grow (Lim 2009: 187). The newly established Singapore adopted a free immigration policy attracting immigrants from Southern China, Malaysia, and India, for instance (ibid.; Schneider 2007: 153). The three most important ethnic groups were the Chinese, who soon became the largest ethnic group controlling the upper end of the social scale, Malays, who were often involved in agriculture, and Indians, many of whom were soldiers of the British army or entrepreneurs (Lim 2009: 187-8). Towards the end of the 20th century, immigration was more closely controlled, so that the largest group of newcomers allowed into the country were Cantonese, especially from Hong Kong (ibid.: 190). Singapore gained independence in 1965, as the attitude towards the British rule changed during WW2 (Schneider 2007: 155).

The fact that the Singapore population consists of a number of different ethnic groups calls for a language that is shared and accepted by the majority of the people. According to Platt (1991: 376), Bazaar Malay had for centuries been used as the lingua franca in the region and held its status during most of the colonial period, but English became a prestige lingua franca among those who knew it. Singapore has four official languages: English is "the language of neutrality and modernity", while Malay, Mandarin and Tamil represent the three largest ethnic groups (Lim 2009: 189). Following the establishment and the success of English-medium schools from the early 19th century onwards, a bilingual education policy

[41] Source: https://www.cia.gov/library/publications/the-world-factbook/geos/rp.html, accessed on September 19th, 2013

requires students to gain proficiency in English and one other language (usually Mandarin, Malay or Tamil). Schneider (2007: 156) presents a number of factors which have contributed to the status of English in Singapore. In addition to the bilingual education policy, the fact that the indigenous languages are mutually unintelligible has made English the binding force of the Singapore people.

Alongside the multilingual situation with four official languages, Singapore English has two major varieties: Standard Singapore English (StSinE) and Colloquial Singapore English (CollSinE). Bao and Hong (2006: 105-6) describe the situation as diglossia in which, in simple terms, CollSinE is used for informal purposes and StSinE on more formal occasions. StSinE differs very little from other standard varieties of English, while CollSinE has distinctive features of syntax and morphology as a result of contact with Chinese and Malay (Gupta 1994: 7). The attitudes of Singaporeans towards the two varieties of English are twofold: on the one hand, there is an official movement promoting the use of StSinE (Speak Good English Movement[42]), and on the other hand, CollSinE is regarded as a marker of Singaporean identity (Wee 2008: 264).

According to Tay (1991: 322), SinE has the widest range of uses of all Southeast Asian Englishes. In addition to being one of the official languages, SinE is the language of education, business and religion. Furthermore, SinE is a language of inter- and intraethnic, as well as international, communication. More significantly, SinE is used to express national identity and is used as a home language, contrary to most other varieties of English used in Southeast Asia.

Table 10. Facts on Singapore.[43]

Number of residents	5.4 million
Significant ethnic groups	Chinese 76.8%, Malay 13.9%, Indian 7.9%, other 1.4% (2000 census)
Languages	Mandarin (official) 35%, English (official) 23%, Malay (official) 14%, Hokkien 11%, Cantonese 6%, Teochew 5%, Tamil (official) 3%, other Chinese dialects 2%, other 1% (2000 census)

[42] For more information, see www.goodenglish.org.sg.

[43] Source: https://www.cia.gov/library/publications/the-world-factbook/geos/sn.html, accessed September 18th, 2013.

Hong Kong. The English language was first brought to China by trading ships in the early 17th century, but more large-scale spread of English into the area happened nearly 200 years later as Western missionaries, merchants and colonial officers were allowed deeper into Chinese territory (Bolton 2002: 4-5, 31). As a result of the First Opium War in the early 1840s, the Hong Kong Island became a crown colony, and stability in the area was guaranteed in 1898, as the New Territories, part of mainland China next to the Hong Kong Island, were leased for 99 years (Schneider 2007: 135). Trade between the British and mainland China was facilitated by Chinese Pidgin English, which emerged in the early 18th century in Guangzhou, and continued to be spoken in Hong Kong until the 1960s (Cheng 1992: 164; Bolton 2002: 5). The use of Chinese Pidgin English declined towards the end of the 19th century as many Chinese learned a more standard form of English from Western missionaries and merchants. Mass education system made English available to people outside the élite in the 1970s, which led to a dramatic rise in the number of English speakers (ibid.: 6). Hong Kong's era as a British colony came to an end in June 1997, as the leasing period of 99 years reached its end. The fact that Hong Kong was handed over to another power instead of gaining independence makes it different from most other postcolonial countries (Schneider 2007: 136).

The latter part of the 20th century saw many changes in Hong Kong society, including its linguistic profile. According to Bolton (2002: 2-3), the 1970s and 1980s witnessed a basically monolingual Hong Kong with Cantonese as the language of the society. Although both English and Chinese (either Cantonese or Putonghua, see Webster 2009: 144) were official languages, the majority of Hong Kong people could not speak English. The Hong Kong government aims at 'trilingualism' and 'bi-literacy', meaning that the Hong Kong people should be able to speak Cantonese, Putonghua and English, and be able to write in Chinese and English (see e.g. Tsui and Bunton 2002: 57). Consequently, in the beginning of the 21st century, approximately 45% of the Hong Kong residents claimed to have knowledge English, while in the 1960s, the proportion was less than 10% (Bolton 2002: 2-3; Webster 2009: 145). While the importance of Putonghua has increased since the Handover in 1997, English is still valued as an international lingua franca. Furthermore, English can be seen as "a social ladder", as the knowledge of English increases the opportunity of securing better jobs, for instance (Webster 2009: 146).

According to Bolton and Graddol (2012: 7), English has fewer uses in Hong Kong than in other postcolonial countries, such as India, Singapore and the Philippines; domains in which English is used in Hong Kong include education,

international business, tourism, and media. The domain of most interest to the present study, personal life, typically uses Cantonese or, especially among university students, code-mixing (Bolton 2002: 11).

The status of Hong Kong English has, until recently, been a source of dispute for both scholars and Hong Kong people themselves. In fact, the term 'Hong Kong English' was suggested by Kingsley Bolton only in the beginning of the 21st century (Schneider 2007: 137). Before that, 'English in Hong Kong' was felt to be a more suitable term for the variety. It is clear now that there are identifiable local features of HKE: vocabulary segments, syntactic features, and a distinct HKE accent (see Schneider 2007: 138 for more information). Although the academic community seems to be accepting HKE as a distinct variety, the Hong Kong society still has its doubts. According to Tsui and Bunton (2002: 58-9), HKE is not accepted by the community that uses it: features of HKE are regarded as deviations from the exonormative model (either BrE or AmE).[44] The government and the business factor have launched several language campaings for the use of 'better' English, similarly to the situation in Singapore (see above).

Table 11. Facts on Hong Kong.[45]

Number of residents	7.1 million
Significant ethnic groups	Chinese 93.6%, Filipino 1.9%, Indonesian 1.9%, other 2.6% (2011 census)
Languages	Cantonese (official) 89.5%, English (official) 3.5%, Putonghua (Mandarin) 1.4%, other Chinese dialects 4%, other 1.6% (2011 census)

Summary. The eight varieties included in the study represent different types of colonization, different lengths of English presence in a given country, different statuses of English, etc. In order to grasp the differences and similarities of the varieties, the varieties were classified according to the three models introduced in

[44] This 'complaint tradition' may be seen as evidence that HKE has reached Phase 3 in Schneider's Dynamic Model.

[45] Source: https://www.cia.gov/library/publications/the-world-factbook/geos/hk.html, accessed on September 19th, 2013.

section 3.2 above (see Table 12 below). The first two models are, in the end, rather similar: the more traditional varieties are labelled ENL or Inner Circle (henceforth IC) varieties, while IndE, PhiE and SinE are ESL or Outer Circle (henceforth OC) varieties. More interesting cases are JamE and HKE: how these varieties should be labelled according to the models is not straightforward.

The linguistic situation of Jamaica raises the question of whether JamE should be considered an IC or an OC variety. The fact that most Jamaicans speak a form of language closer to the basilectal end of the continuum rather than the acrolectal, indicates towards an OC situation: JamE is not the native language of the people, and it is learned in schools and used mainly in writing (see above). Deuber (2014: 67, 248-9) points out that Caribbean varieties of English, including JamE, do not

Table 12. Models of World Englishes.

Variety	ENL/ESL/EFL[46]	Three Circles	Dynamic Model
BrE	ENL	IC	not applicable
IrE	ENL	IC	Phase 5[47]
AmE	ENL	IC	Phase 5
JamE	(ENL/)ESL/ESD	(IC/)OC	Phase 4
IndE	ESL	OC	Phase 3(+)
PhiE	ESL	OC	Phase 3
SinE	ESL	OC	Phase 4
HKE	ESL(/EFL)	OC(/EC)	Phase 3

[46] Only one label is appointed to each country, although the linguistic situation might be more complex and call for a combination of labels. The label is chosen according to the 'mainstream' linguistic situation in the country. The same applies to the 'Three Circles' model.

[47] Ireland is not discussed in Schneider (2007): Phase 5 is assumed to be the appropriate description of IrE.

clearly represent either IC or OC varieties, and may instead "be described as English as a second dialect (ESD)." In an ESD situation, according to Görlach (1991: 12), standard English is considered the 'high', and a pidgin or creole the 'low' variety - this view corresponds with the diglossic perception of the Jamaican linguistic situation. Mair (2011: 211), on the other hand, describes the relationship of JamE, "the emerging local norm of educated English usage in Jamaica", and JamC, a local English-lexicon creole, as a continuum in which most speakers have command of the mesolectal range. Mair (ibid.) further points out that even speakers of the more basilectal end of the continuum consider themselves speakers of *English*. The emerging picture is thus a complicated one: on the one hand, the native language of most Jamaicans is JamC while JamE is mostly acquired through education, and on the other hand, speakers of JamC consider themselves speakers of English rather than a creole. The picture is further complicated by the data used in the present study: Deuber (2009: 1) describes the Jamaican component of the *International Corpus of English* as 'informal JamE' representing the higher end of the post-creole continuum. Whether the face-to-face conversations in ICE-JA truly represent the linguistic situation in Jamaica could thus be questioned, as Mair (2012: 221) does. As regards the classification of JamE as an IC or an OC variety, previous studies have mostly opted for the OC option (see e.g. Collins & Yao 2012, Nesselhauf 2009).[48] In this study, JamE is considered an OC variety, based on the fact that the native language of most Jamaicans is JamC, and JamE is learned in school. This decision is further supported by Mair (2013: 263-4), who situates JamC on a higher level than JamE in his 'World System of Standard and Non-Standard Englishes'[49], and points out that Jamaicans living abroad are more likely to express their ethnic identity through JamC than JamE.

As regards Hong Kong and the variety of English spoken in Hong Kong, the situation is similar, albeit one step farther away from the 'core' of the models in question. As discussed above, the status of HKE has until recently been under dispute, and as the functions of the variety are fewer than of other varieties, it may be questioned whether HKE should be classified as an OC or an EC variety. In this study, however, HKE is considered an OC variety, following Bolton's view of HKE as stabilizing it status in Hong Kong society and in the academic community.

---

[48] According to Deuber (2014: 249), JamE and other Caribbean varieties have often been categorised as IC varieties or ignored altogether: unfortunately, she does not provide examples of studies in which JamE is given an IC status.

[49] JamC is labeled as a 'super-central variety', while JamE is a 'central variety'.

Furthermore, Schneider (2014: 13) presents two studies which indicate that HKE is indeed moving toward the OC, if not already there.

As may be expected, classification of any variety according to any pre-existing model may prove difficult due to the nature of human language: there is no fixed set of phonological or grammatical rules that could be used to strictly define a variety of language; rather, languages and varieties of languages are prone to keep on evolving and changing. In the present study, the classification of the eight varieties has been done based on a number of previous studies. The 'Three Circles' model was chosen to be used in the analysis of the data due to the fact that it is widely used in research on World Englishes. Although Schneider's Dynamic Model offers an interesting point of departure, it is better suited to studies attempting to describe one or more varieties, rather than grammatical features. It seems doubtful that the evolution of the progressive, or any other grammatical feature, could have been tied to the colonial history of a variety of English. Such a connection would suggest that varieties which have reached the same phase would use the progressive in similar ways. Whether or not this is the case will not be discussed in this study.

3.4 Previous research: the progressive in World Englishes

Section 2.2 discussed a number of previous studies on the progressive, excluding those dealing with varieties other than BrE and AmE – such studies are introduced in the present section. Peter Collins' two articles, published in 2008 and 2009, are the ones most similar to the present study regarding the data, methods and eventually also the results. Hundt and Vogel (2011) discuss the progressive in IC, OC and EC varieties, however, their data comprises student writing rather than spontaneous dialogue. Hundt (2009a) focuses on the use of the progressive passive in IC and OC varieties, while Davydova (2012) includes a case study of the progressive in varieties of English spoken in India and Russia. Schilk and Hammel (2014) discuss stative uses of the progressive in South Asian and Southeast Asian varieties of English in order to investigate areal homogeneity. Meriläinen et al. (forthcoming) focus on extended uses of the progressives in IC, OC and EC varieties, thus adding the third circle so far neglected in studies on World Englishes. In the following, three of the studies are introduced in more detail: Collins (2008), Hundt and Vogel (2011) and van Rooy (2014).

Collins (2008). Collins investigates the use of the progressive in altogether nine varieties of English[50], using both written and spoken components of the ICE corpora, and focusing on "the distribution and frequency of progressive forms, meanings and uses" (p. 226). Collins studies the forms that the progressive occurs in, the semantic classes of the main verbs, the uses of the progressive (focusing on the special uses claimed to be involved in the spread of the progressive), and a number of features related to the grammatical environment of the progressive.[51] Collins provides the reader with diverse tables showing the quantitative results of the study and includes a number of different comparisons between various groupings of the varieties. However, at times the discussion of the results is rather brief and cursory. The results of the Collins' study are not presented here, as they are discussed in the relevant sections of the present study. Overall, Collins' findings suggest that the progressive is used differently across the varieties and across geographical regions; the Southern Hemisphere varieties are the ones most different from the others and, in Collins' words, "more advanced, or perhaps 'adventurous'" (p. 247).[52] Collins' study is very similar to this study – the datasets are largely the same, the features investigated likewise – thus a comparison of the results of this study and those obtained by Collins must be conducted carefully.

Hundt and Vogel (2011). In their 2011 study, Hundt and Vogel discuss the frequency and use of the progressive with stative verbs in three groups of varieties representing IC, OC and EC countries.[53] Hundt and Vogel argue in favour of including more than one variety from each circle, as varieties within a single circle have been shown to differ as regards the use of the progressive (e.g. Collins 2008), and furthermore, the varieties included in their study represent different developmental stages, and different first language backgrounds (p. 146). Contrary to expectations set by previous studies on individual varieties or by handbooks on World Englishes, "the study showed that the varieties do not fall into neat groups along the lines of the ENL-ESL-EFL taxonomy" (p. 153). The second feature

[50] The varieties are AusE, NZE, AmE, BrE, KenE, IndE, PhiE, SinE and HKE.

[51] The features are clausal negation, clause type, temporal specification, and contraction.

[52] The results of Collins' study speak for the inclusion of either Australian or New Zealand English in this study, as well, and the lack of an Antipodean variety in the present data is undoubtedly lamentable. See section 4.3 below.

[53] The varieties are BrE, NZE, and IrE; MalaysianE, SinE, PhiE, KenE, and FijiE; and Finnish, Finland-Swedish, and German learners of English. IC and OC varieties are represented by components of the ICE corpora, and EC varieties by components of the *International Corpus of Learner English* (ICLE).

investigated by Hundt and Vogel, combination of the progressive with stative verbs, brings out a difference between the three groups of speakers: IC speakers do not combine the progressive with 'unsuitable' stative verbs[54]; OC speakers (especially Fiji English speakers) do use the progressive (albeit relatively rarely) with stative verbs that BrE speakers would regard unacceptable; while EC speakers, having learned the language in classrooms only, are typically more careful with stative verbs in combination with the progressive (pp. 156-8). According to Hundt and Vogel, the differences between the three groups may be explained by the fact that OC varieties are taking on new contexts where the progressive is used instead of the simple form (p. 158), and on the other hand by the fact that OC speakers use English in creative ways due to being exposed to the language both at school and at home and on the streets (p. 160).

Hundt and Vogel argue that Kachru's 'Three Circles' model may need to be revised due to the fact that distinctions between the three groups, and especially between IC and OC varieties, are blurring (p. 161). The authors call for more research on OC and EC varieties of English, as they may well turn out to be the sources of new global developments (p. 162). As regards the present study, Hundt and Vogel's study could be seen as support for the use of spoken language data: Hundt and Vogel note that the use of the progressive in their written data is rather low and that more varied use of the progressive with stative verbs could be found in spoken data (p. 157). On the other hand, Hundt and Vogel's study suggests that the present study might benefit from adding data from a more recent IC variety, e.g. New Zealand English, or from adding a number of EC varieties to the study. However, as my study was well underway at the time that Hundt and Vogel's study was published, such additions have to be set aside for the time being.

Van Rooy (2014). Van Rooy (2014) focuses on stative uses of the progressive in three Outer Circle varieties of English: Indian, Kenyan and Black South African English. Van Rooy provides a survey of previous studies on stative progressives, criticizing most of them for their unsatisfactory analyses. According to van Rooy, earlier studies regarded the use of stative verbs in the progressive as learner errors (p. 158), while later on researchers acknowledged the existence of this extended use of the progressive. Van Rooy answers the call for more accurate analysis and more elaborate explanations himself, although he does give credit to a few previous studies (e.g. Sharma 2009; Filppula et al. 2009). The aim of van Rooy's paper is

[54] Stative verbs may occur in the progressive resulting in a more dynamic meaning, cf. section 2.1 above.

thus "to determine whether there is mere overextension of the progressive in Outer Circle varieties of English, or whether a different semantic prototype needs to be postulated" (p. 161). The data covers three OC varieties representing both written and spoken material, and is carefully analysed to draw out the stative uses of the progressive.

Van Rooy's results show that the progressive overall is more frequent in spoken rather than written language, and that the African varieties use the progressive more frequently than IndE (p. 164). As regards stative use of the progressive, van Rooy finds that it is more common in the three OC varieties studied than in BrE (van Rooy uses Kranich 2010 as a reference point to BrE). Most importantly, however, van Rooy presents a three-way categorization of stative progressives: *temporary states* are the most prototypical use of stative progressives, found regularly also in IC varieties, whereas *on-going states* and *unlimited states* are more closely connected to OC varieties (pp. 166-7). In ongoing states there is no indication that the state will end soon, but neither is it permanent, while unlimited states indicate that the state goes on for an unlimited amount of time. Van Rooy presents detailed quantitative results on his data, and shows that stative progressives should not be considered learner errors nor 'quirky' use of the progressive, but purposeful extensions of the progressive paradigm.

Chapter 3 has discussed the concept of World Englishes and a number models used to categorize them, as well as introduced the eight varieties included in the study. The following chapter turns the attention to the present study, its methods and data.

4 Methods and data

Chapter 4 introduces the data used in this study, as well as the methods used to identify, retrieve and analyse progressives in the data. Section 4.1 discusses methods related to the four individual components of the study. Section 4.2 introduces the corpora used, and section 4.3 discusses matters related to the retrieval of progressives from the data. Section 4.4 discusses issues related to measuring the frequency of the progressive, and finally, section 4.5 briefly introduces the statistical methods used.

4.1 The components of the study

The progressive is investigated from four perspectives in the present study: Chapter 5 discusses the frequency of the progressive in World Englishes, Chapter 6 investigates a number of morphosyntactic features related to the progressive, Chapter 7 focuses on lexical variation in progressive main verbs, and finally Chapter 8 discusses the functions of the progressive in World Englishes. The methods related to the study of these features are discussed in the following subsections, as well as in the relevant chapters.

4.1.1 Frequency

Frequency of the progressive in World Englishes is of great interest for a number of reasons. First, previous studies (e.g. Mair and Hundt 1995; Kranich 2010) have established that the use of the progressive has increased significantly during the last four hundred years. This study does not include a diachronic dimension, but it is possible to relate the present results to those obtained in previous studies. The fact that this study focuses on spoken data has two implications: the progressive is said to be more frequent in spoken language (see e.g. van Rooy 2014), which leads us to expect that the number of progressives in the present data will be high compared to data including both spoken and written material. On the other hand, as most

previous studies include both spoken and written data, or written data only, the present results do not compare directly to those from previous studies. Nevertheless, the present study contributes to the discussion on the frequency of the progressive by adding the World Englishes dimension: only a few previous studies concentrate on the use of the progressive in varieties other than BrE and AmE (see section 3.4 above). This study seeks to provide answers to questions such as 'Which variety or varieties have the highest and the lowest frequency of the progressive?' and 'Do World Englishes group together somehow regarding the frequency of the progressive?'

In addition to regarding the progressive form as a linguistic feature in itself[55], this study considers the progressive a part of the progressive/non-progressive paradigm. The paradigm is here defined according to form rather than meaning; whether the progressive *semantically* fits a verbal slot was not considered a crucial factor. In other words, the study does not take into account whether it is possible to substitute a non-progressive token with a progressive one. As Smitterberg (2005: 9) points out, there are problems related to the variationist criterion according to which "the referential meaning of a given sentence should not change when the variants are substituted for one another". According to this criterion only such tokens should be considered which do not change the meaning of the sentence regardless of which member of the paradigm is used. However, as it may be difficult even for a native speaker of English, let alone a non-native speaker, to decide whether the criterion is fulfilled, it was decided to include all instances of finite verbs, regardless of whether the substitution of a non-progressive verb phrase with a progressive one, or vice versa, changes the referential meaning of the sentence. The decision is further justified by the fact that the restrictions applying to the use of the progressive (regarding e.g. the type of verbs co-occurring with the progressive) may differ in the eight varieties under investigation; including all finite verb phrases in the paradigm ensures that all possible environments of the progressive are taken into consideration.[56] Section 4.4 discusses issues related to measuring the frequency of the progressive in more detail. Results of the analysis are presented in Chapter 5.

---

[55] Following Smitterberg (2005: 10), when the progressive is studied as a linguistic feature in itself, the focus is on the progressive and factors influencing its occurrence: the non-progressive is not included in this framework.

[56] The only verb phrases excluded from the non-progressive paradigm are set phrases such as *you know* and *I mean*.

4.1.2 Morphosyntactic variation

Chapter 6 focuses on the grammatical environment of the progressive by investigating a number of features which have, in previous studies, been found to be important in relation to the progressive and its use.[57] These features are:

a) Tense (e.g. Is the progressive used in the past tense as often as in the present tense?)

b) Perfect progressives (e.g. How does the progressive combine with the perfect in World Englishes?)

c) Modal auxiliaries (e.g. Are modal progressives used in all varieties of English?)

d) Voice (e.g. Are passive progressives found in spoken World Englishes?)

e) Verb phrase patterns (e.g. What are the patterns in which the progressive occurs in World Englishes?)

f) Other linguistic parameters (e.g. Are progressives located in main or subordinate clauses? How does frequency of contracted forms of BE relate to the frequency of the progressive? What type of temporal adverbials modify progressives most frequently?)

With the information gathered from the study of these features it is possible to draw a picture of the use of the progressive in World Englishes – the forms it takes and the positions in which it can be located. The progressive vs. non-progressive paradigm is discussed in connection to some of the morphosyntactic features, e.g. tense choices and modality.

The morphosyntactic features of each progressive found in the data were coded onto Microsoft Excel spreadsheets. Some features are binary and only require a plus sign to indicate whether the feature is relevant (e.g. passive progressives were marked with '+', while active progressives were left unmarked), while other features are more complex and require a more elaborate coding system (e.g. three types of adverbials). Having completed the coding, the data were placed into tables and figures in the Excel program. Percentages were calculated, and simple statistical testing was performed on the results.[58]

Chapter 6 presents and discusses the results of the morphosyntactic analysis.

---

[57] These previous studies include, for instance, Smitterberg (2005), Römer (2005), and Collins (2008), and were discussed in sections 2.2.2 and 3.3 above.

[58] See section 4.5 for more information on the statistical testing performed on the data.

4.1.3 Lexical variation

The importance of investigating the verbs that appear in the progressive lies in the facts that these verbs may reveal about the progressive itself. According to Smitterberg (2005: 147), "common semantic denominators of verbs that are frequent in progressive verb phrases can give indications of what meanings the progressive tends to co-occur with". In other words, verbs that commonly co-occur with the progressive are considered to share their core semantic denominators with the progressive. In the present study, one of the most interesting factors is whether progressive main verbs are similar throughout the eight varieties studied (discussed in section 7.1). Section 7.2 reports on the results of semantic analysis of progressive main verbs: a random sample of progressive main verbs were categorized into semantic domains following Biber et al. (1999) to investigate what kind of verbs co-occur frequently with the progressive. The present data enables the comparison of progressive and non-progressive main verbs, i.e. the comparison of which verbs tend to collocate with the progressive or the non-progressive. It is also important to study which verbs do *not* appear regularly in the progressive, as differences between varieties in this respect may shed light on recent changes in the progressive – some verbs may have already been accepted to co-occur with the progressive in some varieties, while in the case of stative verbs, for instance, semantic restrictions may apply (discussed in section 7.3). Finally, section 7.4 investigates whether there are verbs that prefer the present or the past tense in combination with the progressive.

4.1.4 Functions of the progressive in World Englishes

The functions in which the progressive is used in present-day English were introduced and discussed in Chapter 2 above. Whether speakers of different varieties of English, spoken the world over, use the progressive in similar functions is the topic of Chapter 8. In order to ensure the data being handled as carefully as possible, the analysis of the functions of the progressive is based on a smaller sample of the present data. Thus, the first 250 instances of the progressive in each corpus were included in the data for the semantic analysis, excluding, however, fragmentary forms and perfect progressives. Fragmentary forms were excluded from the semantic data as it, in some instances, the missing element may be crucial for understanding the function of the progressive (see section 6.5 for more information on fragmentary forms of the progressive). Perfect progressives, on the

other hand, were excluded from the semantic analysis due to difficulties related to the categorization of function.[59] Each of the 2,000 progressives was individually analysed, including enough context to bring out the most probable function. Only one function was assigned to each progressive, although in some cases two or more functions could be considered overlapping.

Although analysis of the functions of the progressive is bound to be subjective rather than objective, it is possible to relate the present results to those obtained by others. For instance, Kranich (2010) notes a significant rise in the proportion of subjective (especially interpretative) progressives in historical data on BrE. Whether recent developments in BrE regarding the functions of the progressive are carried over to other World Englishes is discussed in Chapter 8.

4.2 Choice of data

The choice of varieties to be included in the data was largely data-driven, i.e. such varieties whose representative corpora were available to the School of Language, Translation and Literary Studies at the University of Tampere at the time of compiling data for the study were included. This does not, however, mean that the choices made were arbitrary. Rather, the selection of varieties was governed by a combination of availability and relevance. For instance, ICE-East Africa was, after a closer scrutiny, judged to differ from the other corpora to a great extent and was thus excluded.[60]

BrE and AmE are rather obvious choices as the major varieties and as the historical sources for the postcolonial varieties. Though BrE and AmE are often positioned as reference varieties to which other varieties are compared, the study seeks to regard all eight varieties as equal, regardless of their status as a native or non-native variety. IrE and IndE were chosen because previous research (see e.g. Filppula 2003, 2008; Sharma 2009) has indicated that the progressive is used in both IrE and IndE frequently and in various functions. JamE exists on a continuum with Jamaican Creole, and thus offers an insight into a situation where English is used mainly in the written and more formal spoken modes. PhiE has a linguistic background slightly different from the other postcolonial varieties as the

[59] See Kranich (2010: 180-2) for a discussion on problems related to the combination perfect + progressive.

[60] Schmied (1990) discusses ICE-EA and its compilation.

colony established in the Philippines was American rather than British. SinE and HKE are both situated in a Chinese speaking environment, but their situation today is rather different: Singapore has gained its independence, whereas Hong Kong is now part of the People's Republic of China. The choice of these eight varieties fulfils the requirements of a satisfactory dataset (as discussed by Hundt and Vogel 2011): there are more than one variety from the two Circles; the varieties are in different developmental stages (see Schneider's Dynamic Model); and they represent different linguistic backgrounds with different substrate languages.

These eight varieties of English provide one view of how the progressive is used in World Englishes – a different set of corpora might produce different results. For instance, this study does not include an Antipodean variety of English, although NZE seems to be an interesting variety based on previous studies (see section 3.4 above), nor is an African variety of English included. A third dimension to the study would be gained by including varieties from the Expanding Circle.[61] What this study is able to tell us is how the progressive is used in *these eight varieties*, and the fact that these eight varieties represent IC and OC varieties provides a possibility to discuss how the use of the progressive varies in IC and OC varieties. Whether the varieties included in the study are *typical* representatives of their circle is beside the point; the more important fact is that the Inner and Outer Circles are represented by more than one variety thus avoiding the caveat of having to rely on information from one or two varieties. Hundt and Vogel (2011) stress the importance of including a sufficient number of varieties so that they can be considered to be representative of a larger group of varieties.

The data used in the present study come from the *International Corpus of English* (henceforth ICE) and from the *Santa Barbara Corpus of Spoken American English* (henceforth SBCSAE). Altogether seven components of the ICE are used, and together with SBCSAE they offer a representative view over the use of the progressive in English throughout the world. The focus of the study lies in spoken language, and especially in spontaneous dialogue. Leech et al. (2009: 125) provide the distribution of the progressive in different genres of ICE-GB: the results clearly show that the progressive is most frequently found in categories 'Phone calls', 'Social letters' and 'Conversations (face-to-face)', and least frequently in written academic texts. A simple search for one progressive form (*are* + V*ing*) in all

[61] Meriläinen et al. (fc.) discuss the use of the progressive in IC, OC and EC varieties.

categories of one of the ICE corpora shows that the same preference for casual speech is found in the present data, as well.[62] Thus to restrict the amount of data, the selection of data was limited to a category labeled 'Direct conversations' or 'Face-to-face conversations'. The category in question includes approximately 180,000 words per corpus (see below for more information), but to further limit the amount of data, and to ensure comparability between the corpora, only 100,000 words per corpus were included in the present data.

Subsections 4.2.1 and 4.2.2 below introduce the corpora used in the study, and section 4.3 discusses the retrieval of progressives from the corpus data.

4.2.1 International Corpus of English

The *International Corpus of English* (ICE) has since the early 1990s endeavored to produce comparable corpora representing the many varieties of English spoken the world over.[63] At the time of writing, there are thirteen corpora available, and another eleven are being compiled.[64] Each of the ICE corpora consists of one million words, in the form of 500 separate files of approximately 2,000 words each. The design of the corpora is identical: 60% represents speech and 40% writing, and both public and private domains are included.[65] Although at the moment tagged versions of most of the corpora are available, this study uses the untagged versions, simply because at the time of collecting the present data, the tagging had not been completed.

The ICE corpora represent the speech and writing of adults over 18 years of age, who have been educated through the medium of English, and who were born in or moved into the country they represent at an early age. According to Nelson et al. (2002: 4-5), the main criterion for selecting texts to be included in ICE-GB was the informants' educational level rather than their level of proficiency or the style of language they produced: this was done in order to maintain objectivity in selecting the texts. Nelson et al. (ibid.) refer to the sampling method of the *British* component of ICE, but as the set of corpora is intended to be similar, it may be

[62] See Appendix A for the results of the search.

[63] See Greenbaum (1996) for more information on the compilation of the ICE corpora.

[64] See the project website for the latest information: http://ice-corpora.net/ice/index.htm.

[65] See Appendix B for the design of the ICE corpora.

assumed that the same criterion applies to the other components included in the study, as well.

Table 13. Components of ICE included in the present study.

ICE-Great Britain	ICE-GB
ICE-Ireland	ICE-IRL
ICE-Jamaica	ICE-JA
ICE-India	ICE-IND
ICE-Philippines	ICE-PHI
ICE-Singapore	ICE-SIN
ICE-Hong Kong	ICE-HK

The ICE components included in the study are presented in Table 13 above.

Although the initiators of the ICE project intended the components to be similar and easily comparable, a number of issues to be taken into account have arisen in the course of study. First, ICE-GB is provided together with a search tool of its own, ICECUP. Although ICECUP is a powerful tool, it was decided that all the components should be handled with the same tool, i.e. WordSmith Tools Concordancer[66], to ensure that the components are treated similarly as regards the retrieval of progressives.

Second, the 'Face-to-face conversation' category of ICE-IRL is larger than in the other components, consisting of approximately 210,000 words instead of 180,000 words. However, as the material used in the present study was restricted to 100,000 per corpus, this difference is insignificant.

Finally, and more importantly, some of the corpora contain speech produced by speakers representing some other variety of English, usually by the person recording the conversation (see Nelson et al. 2002: 8-9). For instance, ICE-JA contains speech produced by a German speaker of English, and ICE-HK includes

[66] For more information on WordSmith Tools, see http://www.lexically.net/wordsmith/index.html.

a native speaker of English. The turns uttered by these participants, labeled 'Speaker Z', have been coded differently from turns uttered by speakers representing the variety in question (i.e. HKE in ICE-HK) by the corpus compilers, and it is thus possible to exclude them from the data. Another type of 'extra-corpus material' is text exceeding the length of 2,000 words: the default length of the files included in the corpora is 2,000 words and material exceeding this limit has been tagged to enable excluding it from the data. The extract below shows the beginning of the extra corpus material, marked by <X> in the text and also in the following headers:

(1) <ICE-SIN:S1A-015#241:1:A>

I already wrote to him you know <X>

<ICE-SIN:S1A-015#X242:1:A>

Uh I already told him what's our requirement and he should

prepare that on Thursday

In this study, the extra-corpus material has been excluded from the data. Excluded material includes, for instance, speech produced by Speaker Z in ICE-JA and ICE-HK.[67] Excluding the extra-corpus material, as well as any additional information provided by the compilers of the corpus, is done by setting the Concordancer to ignore text surrounded by angle brackets ('<*>'). As the eight corpora included in the study were handled separately, one after another, the settings had to be set

[67] Unfortunately, there are numerous instances in ICE-HK where the speakers have been coded incorrectly. In the extract below (unnecessary lines omitted here), for instance, speaker A's turn about his experience arriving at an airport is followed by speaker Z's turn after which they both laugh. The following turn (#15 to #28) is marked as speaker Z's but it is quite evident that it is, in fact, speaker A who continues his story. The lines which have been marked incorrectly are not, however, surrounded by <X> and </X>, which means that they have not been excluded from the data. Nevertheless, one should take note of such errors when using ICE-HK.

<ICE-HK:S1A-074#10:1:A>
In fact nobody 's in at the airport waiting for me
<X> <ICE-HK:S1A-074#X13:1:Z>
No one pick you up </X>
<&> <$A> and <$Z> laugh </&>
<ICE-HK:S1A-074#15:1:Z>
That 's that 's my first time to be here

separately for each corpus. However, for one reason or another, the extra-corpus material in ICE-SIN was *not* excluded from the data, i.e. the files from ICE-SIN are included in their entirety, including the words exceeding the 2,000-word limit. As this mistake was noticed at a relatively late stage of the study, and as the results obtained from ICE-SIN have not been affected by the extra-corpus material[68], the decision was made to continue using the data including the extra-corpus material.

In some corpora, e.g. ICE-GB, the text has been cut off when the 2,000-word point was reached, and in some cases, a long conversation may continue in another file. As Table 14 below indicates, the number of files included varies from one corpus to another. This is explained by differences in the policies adopted by the groups compiling the components; the number of files included from ICE-GB (55 files) is larger because the files are, in fact, often shorter than 2,000 words, while the number of files for ICE-HK (41 files) is smaller due to the fact that the files contain more than 2,000 words.[69] The files to be included in the study were chosen randomly: approximately every second file in the directory was included until the total of 100,000 words per corpus was exceeded.[70] The files were included in their entirety rather than cut off at the word limit resulting in slightly more than 100,000 words being included in the data.

[68] The results obtained from ICE-SIN were checked by comparing two sets of data, one including the extra-corpus material and the other excluding it. There were no statistical differences found between the two sets of data, which means that the 51 progressives found in the extra-corpus material are as legitimate as the ones in the main body of the corpus.

[69] The word counts for the files were performed with WordSmith Tools WordList.

[70] See Appendix C for a list of all files included in the present study.

Table 14. Number of files and words from ICE.

Subcorpus	No of files:	No of words:
ICE-GB	55	100,410
ICE-IRL	48	100,028
ICE-JA	47	100,511
ICE-IND	47	101,066
ICE-PHI	46	100,876
ICE-SIN	47	100,812
ICE-HK	41	100,512
Subtotal	331	704,215

4.2.2 Santa Barbara Corpus of Spoken American English

At the time of compiling the data for the present study, the AmE component of ICE was not available.[71] However, a common practice has been to use the *Santa Barbara Corpus of Spoken American English* (SBCSAE) to compensate for the missing spoken AmE component (see e.g. Collins 2008: 228). SBCSAE consists of approximately 249,000 words of naturally occurring speech recorded across the United States of America, and the informants represent different ages, genders, occupations, religions, and ethnic groups.[72] Most of the recordings represent face-to-face conversations, but also other spheres of everyday communication, e.g. telephone conversations or game-playing, are included. To ensure compatibility with data drawn from the ICE corpora, the files included in the present data were

[71] At the time writing, the written component of ICE-USA is downloadable from the ICE website.

[72] See the project web page at http://www.linguistics.ucsb.edu/research/santa-barbara-corpus for more information.

Table 15. Number of files and words from SBCSAE, and in total.

Subcorpus	Number of files:	No of words:
ICE	331	704,215
SBCSAE	21[73]	100,008
Grand total	352	804,223

selected from the face-to-face conversation section only. Table 15 above shows the number of files and words in SBCSAE and in the data as a whole. The format of the SBCSAE files is not identical to those of the ICE files, however, it is possible to use the same methods to retrieve and analyse progressives in the two sets of corpora.

4.3 Retrieval of progressives

Having chosen the varieties to be included in the study, and having considered the delimitation of the corpus data to be used (i.e. 100,000 words of casual conversation), it is now time to turn the attention to the progressive itself. In the study, the progressive is considered to be formed with a form of the verb BE and the present participle of the main verb, i.e. BE + V*ing*. The progressive aspect may be indicated by other means as well (e.g. in JamC and HKE[74]), but these other possibilities of expressing the progressive are not included in the present study. Also, alternative spelling forms are possible for the present participle (*-in'*, *-in*), but the orthography of the ICE corpora uses the fully spelled form. Thus to retrieve progressives from the data, a number of search strings spanning all the possible forms of the verb BE followed by a word ending in *-ing* were entered into the WordSmith Tools Concordancer. As is it possible that the auxiliary verb BE and the

[73] As the size of files in SBCSAE is not limited to approximately 2,000 words as in ICE, the number of files is considerably lower than for the ICE corpora.

[74] Sebba (1997, cited in Deuber 2014: 5) discusses a continuum of expressing the progressive in Jamaica: at the one end, the StE *I am eating* is used, and at the other end we find the creole progressive *mi a nyam*. Variants towards the lower end of the continuum are not included in this study. For alternative progressive constructions in HKE, see section 6.5 below.

–ing-form are separated by one or more words, the search was refined so that the form of BE should be located within five words to the left of **ing*. The searches were launched separately on each of the eight corpora,

Table 16. Search strings for the progressive.

Context, 5 words to the left	Search word	Retrieves
am/'m/are/'re/aren't/ is/'s/isn't	**ing*	present progressives
was/wasn't/were/weren't	**ing*	past progressives
can/could*/may/might/will/*'ll/would*/ must*/shall/should**	*be *ing*	modal aux. + progressives
has/hasn't/have/haven't	*been *ing*	perfect progressives
Ø	*being*	passive progressives[75]
(Context excluding all of the above)	**ing*	fragmentary progressives

and the results of the search were also handled separately for each corpus. Table 16 above portrays the searches used to retrieve progressives from the data.

Retrieving all instances of **ing* preceded by any form of BE within five words produced a large amount of unwanted hits, either because the word ending in *–ing* was not a present participle (but a noun, an adjective, etc.: see (4.2) and (4.3) below) or because one of the pre-designed restrictions applied to it. Following previous research on the progressive, Smitterberg (2005) in particular, a number of restrictions were taken into account during the manual checking of the tokens produced by the corpus searches. Instances excluded from the data include *to*-infinitives (as in (4.4)), BE *going to* with future reference (as in (4.5)), appositively

[75] Search string "Ø *being*" retrieves passive progressives, but also progressives of the verb BE. The output of this search (and any other search) was manually checked to ensure that only relevant hits were included.

used participles (as in (4.6)), BE + gerund (as in (4.7)), and repeated or unclear tokens (as in (4.8) and (4.9)).

(4.2) But it 's a big thing (ICE-HK:S1A-063#298:1:A)

(4.3) It 's terribly exciting (ICE-GB:S1A-049 #332:1:A)

(4.4) But I 'm not supposed to be saying (ICE-GB:S1A-017 #337:1:C)

(4.5) I think he's going to drive us to the disco (ICE-SIN:S1A-009#145:1:A)

(4.6) people are out partying (ICE-PHI:S1A-007#124:1:A)

(4.7) That is that is playing fouls (ICE-IND:S1A-080#214:1:B)

(4.8) I 'm just kidding I 'm just kidding[76] (ICE-PHI:S1A-017#186:1:A)

(4.9) I wasn't there but I feel<?>ing </?> sorry (ICE-JA:S1A-083)

Although some previous studies do include progressives with *to*-infinitives in their data (e.g. Collins 2008), they were excluded from this study to limit the amount of data and also to ensure the homogeneity of the progressives retrieved from the data. BE *going to* + infinitive with future reference is included by some studies and excluded by others according to how the construction is defined; the more recent studies seem to exclude rather than include it.[77] Although the construction could be regarded as a progressive according to its form (progressive form of the verb GO followed by a *to*-infinitive), Smitterberg (2005: 33) points out that there is no relationship between the progressive and the non-progressive of the verb GO.[78] Mair and Hundt (1995: 114) show that the frequency of the BE *going to* –future has increased during the late 20th century, while at the same time, the frequency of other progressive instances of GO has decreased. Thus in the present data, instances of *going* were included in the data only when there was an "obvious reference to already ongoing motion" (Kranich 2010: 15).

[76] The first occurrence of *kidding* was included in the data, whereas the second one was not.

[77] E.g. Mair and Hundt 1995, Smitterberg 2005, Collins 2008, and Kranich 2010 all exclude BE *going to* + inf.

[78] Smitterberg (2005: 33) gives the following examples to illustrate the difference between the progressive and the non-progressive of GO, and points out that there is no sense of movement related to the progressive in (a):

(a) I am going to sell my house. (future reference)

(b) I go to sell my house.

With regard to appositively used participles, the locative *out* in (4.6) above indicates that *partying* is, in fact, in apposition to rather than part of a progressive verb phrase. Similarly, instances of **ing* preceded by an existential *there* were excluded from the data. The word ending in *-ing* in a BE + gerund construction (see (4.7) above), on the other hand, is regarded as noun-like rather than a present participle and thus excluded.

Finally, tokens repeated by the same speaker were counted only once, as in example (4.8) above. However, if a progressive was repeated, or echoed, by another speaker, both occurrences were included. Unclear tokens were excluded from the data for obvious reasons. Instances such as (4.9) where the transcriber has difficulties in interpreting what they hear cannot be included in the data.

After the unwanted tokens had been manually removed from the results of the corpus searches, the progressives were extracted to eight separate Microsoft Excel spreadsheets, one for each variety. Each progressive was given its own entry, i.e. progressives sharing the form of BE were each placed on their own row. The Excel spreadsheets functioned as the hotspot for the study: most of the linguistic and extralinguistic analyses were coded onto the spreadsheets, after which calculations on each feature were carried out, with the tools provided by Microsoft Excel.

4.4 The progressive vs. the non-progressive[79]

It has been argued (see e.g. Smitterberg 2005, Williams 2002) that it is not sufficient to investigate the progressive alone and compare the number of progressive tokens to the number of words in the corpus. Instead, one should take into account the fact that the progressive form can only occur with verbs: a noun cannot take the progressive aspect, nor can an adverb. Thus, the comparison of the number of progressives should be made to the number of verbs in the corpus. Smitterberg (2005: 37-53) discusses this methodological issue in detail. In brief, Smitterberg summarises the benefits and drawbacks of the three most common ways of measuring the frequency of the progressive, and goes on to suggest a new, fourth solution to the matter, the S-coefficient (2005: 45). The most common way of measuring the frequency of the progressive is the so-called M-coefficient (M

[79] This section is largely based on section 3.1 in Suoniemi (2012).

after Mossé), where the number of progressives is related to the number of words in a text, i.e.

$$M = N_{PROG} / N_{WORD} \times 100{,}000$$

According to Smitterberg (ibid.: 40-2), the M-coefficient has certain advantages that make it a powerful tool: it is easy to calculate, it is objective, and it compares easily to results from other research. However, the M-coefficient fails to take into account the fact that the progressive cannot occur in any slot in a text, which may lead to problems when comparing different genres, for instance. Furthermore, problems may arise when variation between texts is being tested for statistical significance. Nevertheless, the M-coefficient remains the most common device for measuring the frequency of the progressive.

A coefficient that takes into account the restrictions on the use of the progressive is the K-coefficient, i.e.

$$K = N_{PROG} / (N_{VERB} - N_{NOPROG}) \times 10{,}000$$

Verb phrases where the progressive cannot occur are excluded altogether. In other words, "the K-coefficient gives the researcher an indication of the probability of a progressive occurring in a text" (p. 43). Calculating the K-coefficient requires the researcher to judge whether or not a verb phrase could appear in the progressive, which makes it a subjective method, as well as arduous.

Smitterberg reports on a third coefficient which, according to him, has the advantages of the K-coefficient, but avoids some of the problems related to both the M- and the K-coefficients. The V-coefficient simply relates the number of progressives to the number of finite verb phrases in a given text, i.e.

$$V = N_{PROG} / N_{VERB} \times 10{,}000$$

The V-coefficient is fairly easy to calculate, even if one is working with a non-tagged corpus of moderate size. There are no subjective elements to it, and furthermore, it neutralizes the differences between the ratio of the number of verb phrases to the number of words.

Finally, Smitterberg goes on to suggest a fourth coefficient, specially designed for the study of late Modern English (pp. 45-8). The idea is to exclude from the calculation all such *contexts* where the progressive does not occur. For instance, the

formula Smitterberg devices for the late Modern English excludes imperative verb phrases and BE *going to* + infinitive constructions with future reference. With modifications that take into account contexts where the progressive does not occur, the so-called S-coefficient could also be used for other periods of time. However, this would require a great deal of work as there are no clear answers to where exactly the progressive is possible or not in the varieties studied here. Furthermore, restricting the data by excluding a context could skew the results as it is possible that a variety made use of the progressive in a context judged 'erroneous' by native speakers.

For this study, two of the coefficients discussed above were chosen, the M-coefficient and the V-coefficient. The M-coefficient is used for comparisons to previous research, while the V-coefficient is used for the comparison between the varieties included in the study. This applies to the discussion on the frequency of the progressive only; Chapters 6 and 7 make use of raw frequencies and percentages.

Calculating the M-coefficient presents no problems as all one needs to know is the number of progressives and the number of words in a text. The number of words in a text, in this context, refers to the number of words in each corpus sample, i.e. approximately 100,000 words per corpus. The V-coefficient, in turn, required considerably more work, as the study uses the untagged versions of the corpora. In practice, this meant that all verb phrases were located manually in the corpora, and tagged with a set of codes representing four parameters of interest (see Table 17 below, and Appendix D for all possible tags and examples). The progressives retrieved on the first round using the untagged corpora were cross-checked against the tagged data: some progressives had actually been missed on the first round and were thus added to the results, and some hits turned out to be unwanted, after all. Although manual handling of the data is arduous, it ensures that each possible progressive is checked individually.

Table 17. Tags for encoding verb phrases.

Tag	Stands for
P / V	Progressive / Non-progressive
R / A	Present tense / past tense
M (where relevant)	Modal auxiliary included
F (where relevant)	Future reference

Sentences (4.10 to 4.12) below exemplify the tags; _VR refers to a present non-progressive verb phrase, _PA to a past progressive verb phrase, _PRM to a present progressive with a modal auxiliary included, and finally _PRF to a progressive with future reference (indicated in the context).

(4.10) Here I think_VR this is_VR really weird <#> Was I telling_PA somebody this last night in the pub <#> (ICE-IRL, S1A-006)

(4.11) You must have been hurting_PRM this last season. (SBCSAE, sbc0056)

(4.12) It's_VR about uh sixty-five hours but I'm going_PRF via Hyderabad because I'm leaving_PRF on the twenty-fifth evening (ICE-IND, S1A-034)

Tagging the verb phrases with codes related to tense, modal auxiliaries and reference to future events enables a number of comparisons between the progressive and the non-progressive. For instance, it is possible to compare the use of present tense in these two variants of the verb phrase, or to discuss modality in progressive and non-progressive verb phrases.

To sum up, the data in the present study were subjected to two kinds of tools for measuring the *frequency* of the progressive. The M-coefficient

$$M = N_{PROG} / N_{WORD} \times 100{,}000$$

is used for comparisons to results obtained in a number of previous studies, and the V-coefficient

$$V = N_{PROG} / N_{VERB} \times 10{,}000$$

is used for comparison between the varieties included in this study. For all other features, discussed in Chapters 6 and 7, raw frequencies and percentages are used. This applies also to comparisons to previous studies, as the present data is restricted to 100,000 words to begin with. The actual size of data per corpus is close enough to 100,000 to enable crude comparisons to results obtained in previous studies using the M-coefficient. Otherwise, the comparisons are based on percentages.

4.5 Statistical conventions

The results of the study are subjected to the chi-square test to determine whether the differences found in the data are statistically significant.[80] The level of significance is signalled with three options: 0.05 (indicated by *), 0.01 (indicated by **), and 0.001 (indicated by ***). Throughout the tables presented study, asterisks are used to show the level of statistical significance. The chi-square test is only applied to the data if the expected frequency of any cell is above 5; if it is possible to avoid low frequencies by re-grouping the data, such re-grouping is done by combining appropriate categories. Low frequencies are expected to occur when discussing a low-frequency feature such as the progressive passive.

All of the tests performed are based on the tables included and discussed in the present study. The chi-square tests are performed on two varieties at a time, as subjecting all eight varieties to the test as a whole is likely to produce a non-random relation automatically. This is due to the fact that a sufficiently large amount of data will always produce a result that the relation is non-random (Kilgarriff 2005: 263f.). Because of this, the chi-square tests are performed on two varieties at a time: the varieties are ordered according to the frequency of the feature in focus, and the statistical testing is performed on two varieties located next to each other. For instance, the frequency of feature X is the highest in variety A and second highest in variety B, and so forth; statistical testing is then performed on varieties A+B, B+C, C+D, etc.

Appendix E provides the data and the results of the chi-square tests performed.

[80] The chi-square tests were performed with the R program. For more information on R, see the website: http://www.r-project.org/.

5 Frequency of the progressive

5.1 On the spread of the progressive

The frequency of the progressive has been the focus of numerous previous studies, including Strang (1982) on 18th to early 20th –century fiction, Elsness (1994) on Old to early Modern English, Arnaud (1998) and Smitterberg (2005) on 19th-century BrE, Smith (2005) and Leech et al. (2009) on late 20th-century BrE and AmE, and Kranich (2010) on the frequency of the progressive in Modern English. Most of the studies focus on BrE usage of the progressive, or compare the frequency of the construction in BrE and AmE in the 20th century. What all of the studies mentioned indicate is a continuing rise in the use of the progressive, starting from the 16th century (Elsness 1994) and continuing well into the present day (Kranich 2010). Figure 3 below presents the increase in the frequency of the progressive from the first half of the 17th century to the present day (second half of the 20th century).[81] The S-shaped curve represents the typical progress of language change (see e.g. Denison 2003), in which an innovation is taken into use gradually, with a rapid rise as more and more acceptance is gained, and finally the rise is leveled out as the construction has spread into various contexts with no more places to go to. Kranich (2010: 107fn.) points out that the rise of the progressive follows the typical S-shaped curve, and that the final stages of spread have possibly not been reached yet, as the final time span still shows a relative rise instead of leveling out into a plateau. The frequency of the progressive exhibits a gradual rise starting from the 17th century, while the 19th century clearly shows a rapid rise which seems to continue to the first half of the 20th century. In the latter half of the 20th century, the era of rapid rise decreases to some extent, but still shows increasing use of the construction.

[81] Figure 3 is reproduced based on information in Table 3b in Kranich (2010, 95). Note that the results in Figure 3 are based on written genres in the ARCHER2 corpus, and are thus not directly comparable to results in this study.

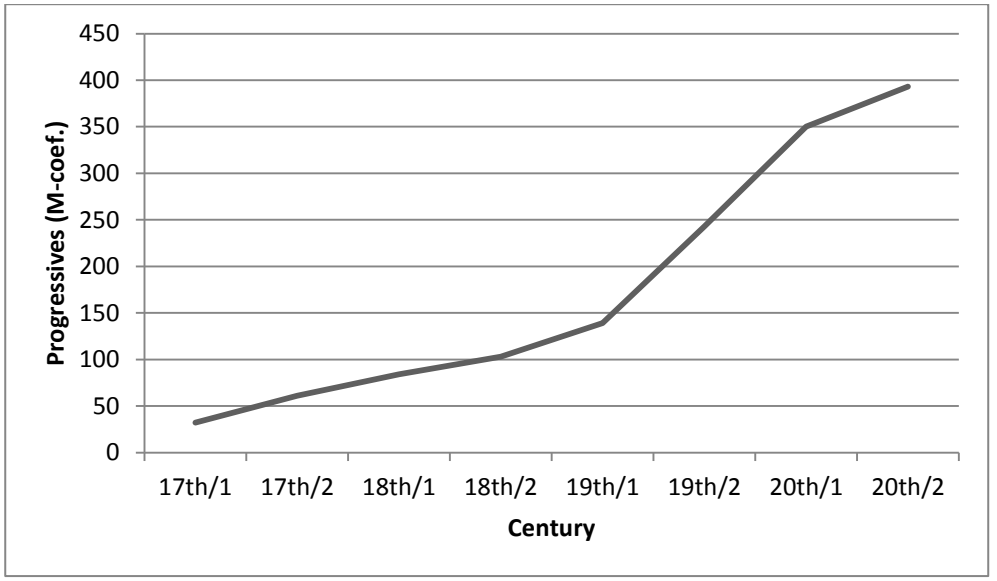

Figure 3. Rise of the progressive in Modern English (based on Kranich 2010: 95).

While most studies on the spread of the progressive focus on written data, there are a few that shed light on what is happening in speech. Such studies include Smith (2005) and Leech et al. (2009), which both compare the use of the progressive in BrE and AmE in the 1960s and the 1990s. Smith's primary interest is on written language, but his brief discussion on the spoken data reveals an interesting fact about face-to-face conversation, i.e. the focus of this study. Smith's data[82] shows that contrary to most written genres and to all other spoken genres included in his study, the frequency of the progressive in face-to-face conversations actually remains stable over the 30 years under scrutiny. Smith suggests that, in conversational use, the frequency of the progressive had already "reached a plateau" by the 1960s (ibid.: 76). Kranich's (2010) data could be read as confirmation for this explanation, as the S-curve portrayed in Figure 3 above starts to level out during the early 20th century. Interestingly, Leech et al. (2009), using samples from an ultimately similar set of corpora[83] as Smith (2005) and focusing on

[82] Smith's primary data consists of the Brown family of corpora (representing written BrE and AmE in the 1960s and the 1990s), but the spoken data consists of samples of 80,000 words each from the *Survey of English Dialects* and ICE-GB (Smith 2005: 50).

[83] The spoken data in Leech et al. (2009) comes from *The Diachronic Corpus of Present-Day Spoken English*. The 1960s data in DCPSE is based on the London-Lund Corpus, while the 1990s data

the same period of time, the 1960s and 1990s, reach a result showing steep increase in the frequency of the progressive also in face-to-face conversations. What has lead to this discrepancy in the results of two very similar studies could be small differences in the selection of texts in the sample corpora or differences in the retrieval of progressives. As the present study is not designed to shed light on the diachronic changes in the spoken usage of the progressive, it suffices to say that, on the whole, the spread of the progressive has been shown to have taken place in written data and also in spoken data, although the results obtained from previous studies are indecisive on what exactly has happened in face-to-face conversations.

Most of the previous studies on the progressive focus on BrE or on a comparison between BrE and AmE. The present study, however, seeks to investigate the use of the progressive in altogether eight varieties of English throughout the world. There are only a few similar studies available so far: two of these are Collins (2008), and more recently, Hundt and Vogel (2011) (see section 3.4 above). Collins' (2008) study focuses on four IC and five OC varieties and shows that there is variation within individual varieties, but not so much between the two groups of varieties (2008: 229). Figure 4 below is reproduced from information presented in Collins' Table 1a, showing the frequency of the progressive in spoken data only.[84] Based on Collins' results, the progressive is most frequently found in the Kenyan component of the ICE corpora[85], followed by the two Antipodean varieties, AusE and NZE, together with HKE. Figure 4 also shows that the differences between the individual varieties are not great, with the exception of KenE versus any other variety.

comes from ICE-GB (similarly to Smith's data) (Leech et al. 2009: 44.) For more information on DCPSE, see http://www.ucl.ac.uk/english-usage/projects/dcpse/.

[84] Collins (2008) presents results for writing and speech separately. Also, Collins presents only raw frequencies for his data; here the data has been normalized to 100,000 words per corpus (Collins' raw data includes 60,000 words of speech per corpus).

[85] Kenyan English forms a part of ICE-East Africa, together with English spoken in Tanzania.

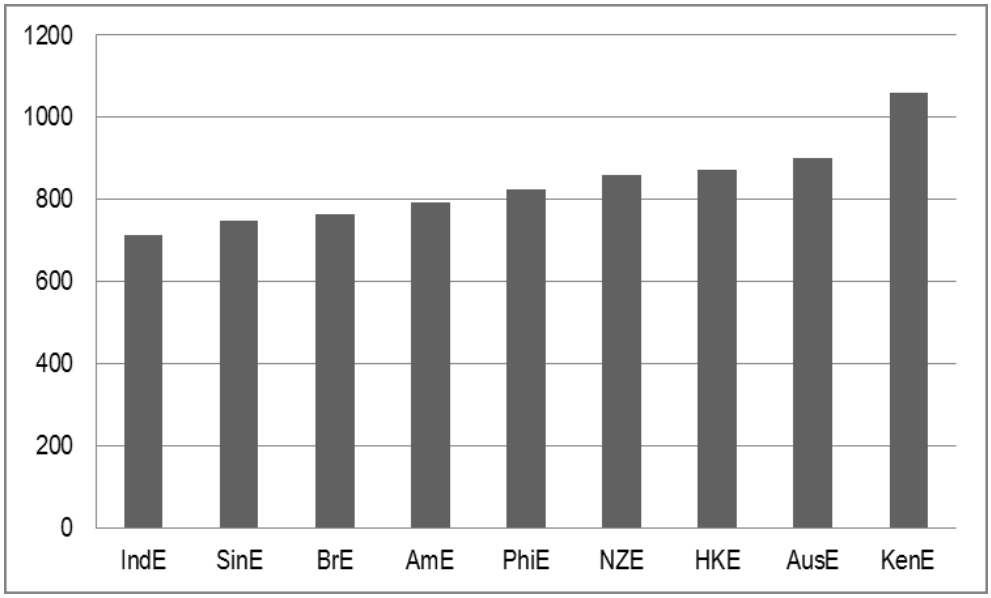

Figure 4. Frequency of the progressive in spoken data (based on Collins 2008: 229).

There is a growing interest in the interplay of IC, OC and Expanding Circle varieties of English that extends to the study of the progressive, as well. Hundt and Vogel (2011) include in their study three IC, five OC and four EC varieties. Their results on the frequency of the progressive are presented in Figure 5 below.[86] Similarly to Collins' results, differences in frequency are found between individual varieties rather than between the three groups of varieties. Most of the ICLE learner corpora do use the progressive more frequently than most IC or OC varieties, but there is the exception of ICLE-Finnish being situated close to the OC varieties, and ICE-Fiji (an OC variety) holding the number one spot on the right. The two traditional IC varieties, BrE and IrE, use the progressive relatively infrequently, with only ICE-SIN showing lower frequency of the progressive.

[86] The frequencies presented by Hundt and Vogel have been normalised per 1 million words (on the basis of information provided in Table 1-a in the Appendix, p. 165) and represent student writing (essays and exam scripts). ICLE-German1 and ICLE-German2 refer to different versions of the German component of the *International Corpus of Learner English*.

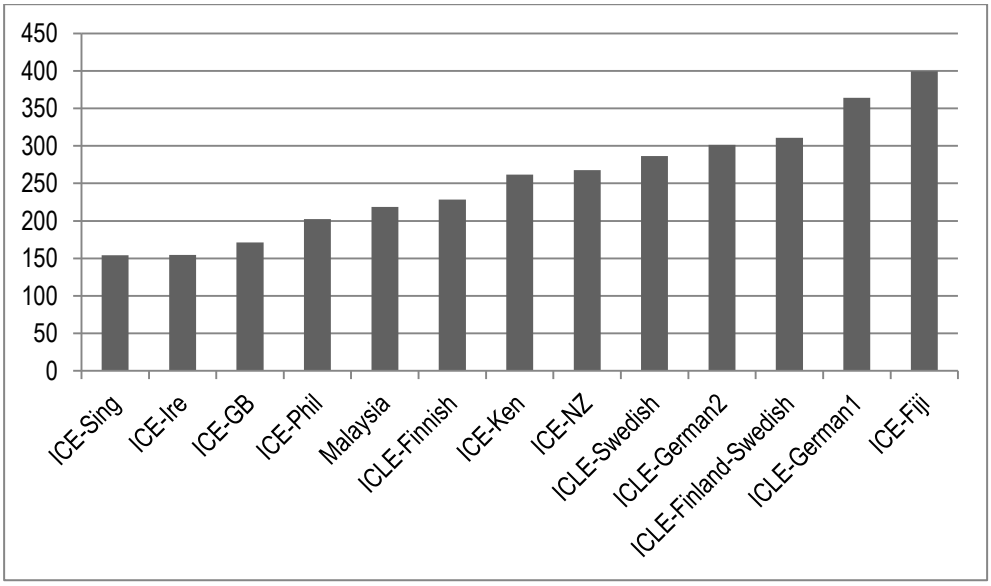

Figure 5. Frequency of the progressive (based on Hundt & Vogel 2011: 154)

These two previous studies on the progressive in World Englishes give reason to make a number of predictions related to the present study. First, differences may be expected to arise between individual varieties rather than between groups of varieties. Second, SinE may be expected to have a low frequency of the progressive in this study, as well, and be located close to BrE and/or AmE on the scale of varieties preferring the use of the progressive. And third, PhiE is likely to appear somewhere in the middle on the scale. Although such predictions are made, it must be borne in mind that the varieties included in the three studies differ, and more importantly, that only Collins' study represents spoken language usage, similarly to the study at hand.

In relation to the increasing use of the progressive throughout the past centuries, scholars have discussed possible reasons for the development of the progressive. One of the most often mentioned reasons is the extension of the progressive to stative situations. In prescriptive literature on StE, the use of the progressive with stative verbs is often considered unacceptable or, at least, non-standard (see e.g. Leech and Svartvik 1994: 74), and only recently are stative progressives gaining acceptance and recognition. Two rather exhaustive lists for possible reasons for the increasing use of the progressive are given by Mair (2006) and Smith (2005). Mair (2006: 88-89) offers three possible reasons for the rise of the progressive. These are:

a) the increasing frequency of many already established uses of the progressive

b) the creation of new forms to fill in gaps in the progressive paradigm

c) the greater tendency to use the progressive with stative verbs.

These points are further discussed in the light of results from the present study in section 6.5 (regarding b)) and Chapter 8 (regarding a) and c)).

Smith (2005: 234-237) also offers a number of possible factors contributing to the increasing use of the progressive. These are:

d) grammaticalization (spread of the progressive to uses and functions more abstract than the original, or more subjective than the original)

e) colloquialization (a drift towards the norms of spoken rather than written language)

f) language contact (especially the influence of AmE on BrE, as Smith's study focuses on these two major varieties)

g) influence of northern and/or Celtic dialects (the progressive is considered more frequent in Celtic varieties of English).

Points d) to g) are further discussed in the light of results from the present study in section 6.5 and Chapter 8 (regarding d)), section 6.6.2 (regarding e)) and section 5.3 (regarding f) and g)).

The frequency of the progressive in World Englishes is presented in the following section.

5.2 Frequency of the progressive in World Englishes

As discussed in subsection 4.4, the frequency of the progressive is measured in two ways in this study. The more traditional M-coefficient (i.e. normalization to 100,000 words) is used for comparison to previous studies, and the more recent V-coefficient (i.e. relating the number of progressives to the number of all tensed verb phrases in the corpus) is used to discuss variation within the varieties included in this study. The present section is thus divided into two: first, the frequency of the progressive in World Englishes is measured and compared to results from previous studies using the M-coefficient, and second, the frequency of the

progressive in World Englishes is measured using the V-coefficient. There follows also a brief methodological discussion related to the use of the two different types of measurement. The next section (5.3) discusses variation in the frequency of the progressive within the varieties included in the present study in more detail.

Table 18 below presents the frequency of the progressive in World Englishes, normalized to 100,000 words per corpus. As the corpus size has been restricted to 100,000 as closely as possible, there are no great differences between the raw frequencies and the normalized frequencies.[87]

Table 18. Frequency of the progressive in World Englishes, M-coefficient.

Variety	Words/corpus	Progressives (RF)	M-coefficient
BrE	100,410	842	839
IrE	100,028	1073	1073
AmE	100,008	771	771
JamE	100,511	859	855
IndE	101,066	890	881
PhiE	100,876	802	795
SinE	100,812	733	727
HKE	100,512	375	373
TOTAL	804,223	6,345	789

The distribution of progressives in the present data is more clearly seen in Figure 6 below which presents the information gathered in Table 18 in a different form, i.e. ordered according to the frequency of the progressive.

[87] The results presented in Table 18 confirm that the M-coefficient is very close to the raw frequency of the progressive in the present data, thus there is no need to calculate the M-coefficient for the morphosyntactic, lexical, and semantic features investigated in the present study.

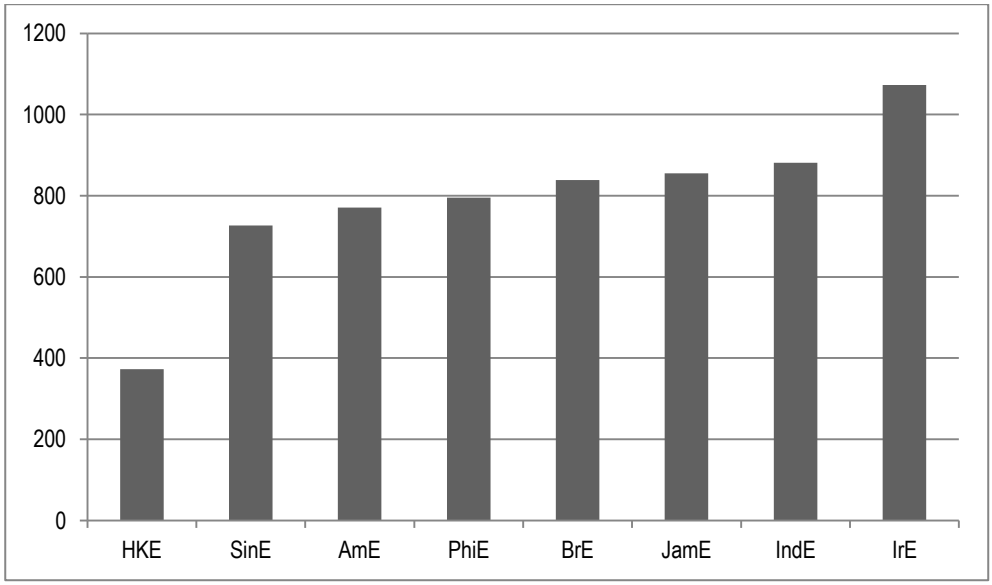

Figure 6. Frequency of the progressive in World Englishes, M-coefficient.

Focusing on differences between results from this study and those of Collins' (2008), it is evident that, on the whole, there are very few of them, as expected. Collins' data is smaller (60,000 words per corpus), and as his choice of varieties is slightly different, results from altogether six varieties can be compared (excluding IrE and JamE). The only major difference in the retrieval of progressives from the data is the fact that Collins includes *to*-infinitives collocating with the progressive, contrary to the present study. Thus the present results can be expected to be slightly lower than those obtained by Collins, and in many cases, this is true. The average frequency of the progressive in World Englishes is M=837 in Collins' data, and M=789 in the present data. The higher frequency of the progressive overall in Collins' data may be partly due to the selection of varieties, e.g. the high frequency of the progressive in KenE. As regards AmE, PhiE and SinE, the differences in the results between the two studies may be due to the small difference in the retrieval methods (i.e. the inclusion of *to*-infinitives). Regarding BrE, this study has a slightly higher frequency for progressives than Collins, thus flipping the order of the two major varieties, BrE and AmE. A number of previous studies have found that the progressive is used more often in AmE than in BrE (see e.g. Smith 2005, Biber et al. 1999), while Mair and Hundt (1995) and Leech et al. (2009) report that the frequency of the progressive is higher in BrE than AmE. The most plausible explanation for the difference between the present study and the previous research

is that of genre: most previous studies focus on written language, whereas the present results come from spoken language. This does not, however, explain the difference related to Collins (2008). Furthermore, the suitability of the Santa Barbara corpus as a replacement for the missing spoken ICE-US has been questioned (see e.g. Suoniemi 2012: 214fn.), but if this was the case, it should be evident in Collins' data, as well.[88]

Another difference found regarding the frequency of the progressive in this study and Collins (2008) is the result for IndE. Collins' data shows a low frequency (M=713) for the progressive, in fact the lowest of all varieties, whereas in this study, the frequency of the progressive in IndE is the second highest of all varieties (M=881). Balasubramanian (2009) and Davydova (2012) argue that the frequent use of the progressive is one of the most typical features related to IndE, which would support the finding from the study at hand.

The most puzzling difference between this study and Collins (2008) is, however, found in relation to HKE. Collins' (2008) results place HKE very high on the scale of frequencies of the progressive: of the varieties shared in the two studies, Collins' results place HKE as the variety using the progressive most often (M=872). In contrast, the present study places progressive usage in HKE at the very low end of the scale, separated clearly from the other varieties (M=373). Such a difference must be related to a fundamental difference in how the data from ICE-HK was handled. Unfortunately, Collins does not discuss how progressives were retrieved from individual corpora, i.e. whether he took into consideration the small differences between the individual components of ICE. The key factor regarding ICE-HK could be the fact that it includes rather a large amount of native speaker data, which in this study has been excluded from the searches (see section 4.3).[89] Indeed, one of the examples given by Collins comes from speaker Z (example 14, taken from ICE-HK: S1A-003), which indicates strongly that he has included all data available in the files without considering the fact that a large part of it comes from either a native speaker of English or a native speaker of a language other than

[88] Furthermore, a similar result placing the frequency of the progressive higher in BrE than AmE was obtained from a query into the American National Corpus, yielding M=650 progressives.

[89] If data produced by Speaker Z in ICE-HK was included in the present study, the frequency of the progressive in HKE would be M=761, which would put the frequency of the progressive higher in HKE than in SinE.

Cantonese.[90] Thus the results regarding HKE in the two studies cannot be compared.

It is rather difficult to find other studies that the present results could be compared to (see section 3.4 above). Hundt and Vogel (2011) share only four of the varieties studied here, and furthermore, their data comes from a specific written genre, i.e. student writings, rather than face-to-face conversations. What can be said, however, is that the normalized frequencies for the progressive in Hundt and Vogel's (2011) study are considerably lower than in the present study, confirming the fact that the progressive is indeed more frequently used in spontaneous speech than in task-based writing in all four varieties shared in the two studies. Besides Collins (2008) and Hundt and Vogel (2011) there are few large-scale comparative studies of the use of the progressive in World Englishes. Meriläinen et al. (fc.) discuss IC, OC and EC varieties of English as regards the extended uses of the progressive in speech. Their results show that the progressive is least frequently used in EC varieties, and that it is most frequent in KenE. Further comparison to the results presented in Meriläinen et al. is futile, as it shares some of the data with this study.

The discussion so far has been based on the M-coefficients calculated for the frequency of the progressive in World Englishes. Table 19 below presents the same raw frequencies of the progressive, now related to the number of verb phrases in the data rather than 100,000 words of any kind. As discussed earlier in subsection 4.4, relating the number of progressives to the number of verb phrases gives a more accurate picture of the use of the progressive, as the verb phrase is the only position in a sentence in which the progressive can occur.

[90] The speakers included in ICE-HK have Cantonese as their first language (cf. Bolt & Bolton 1996: 199).

Table 19. Frequency of the progressive in World Englishes, V-coefficient.

Variety	Verb phrases/corpus	Progressive VPs (RF)	V-coefficient
BrE	14,570	842	578
IrE	13,553	1073	792
AmE	14,348	771	537
JamE	12,670	859	678
IndE	12,238	890	727
PhiE	13,039	802	615
SinE	13,155	733	557
HKE	10,416	375	360
TOTAL	103,989	6,345	610

There are approximately 13,000 verb phrases per 100,000 words in each of the corpora studied. The number of verb phrases is clearly higher in BrE and AmE than in the other corpora (more than 14,300 VPs), which results in the fact that the V-coefficient for BrE and AmE is relatively lower. Indeed, if compared to the results obtained by the M-coefficient, BrE and AmE have shifted a notch lower on the scale of frequency of progressives. In contrast, the number of verb phrases is considerably lower in HKE than in the other varieties (10,400 VPs), which could affect the position of the variety on the same scale. However, as the number of progressives in HKE is so much lower than in the other varieties, there is no effect on the order of varieties. Figure 7 below shows the varieties in the order of most frequent use of progressive on the right. If contrasted with Figure 6 above, Figure 7 shows a more subtle cline from frequent to less frequent use of the progressive, with HKE as the only variety to stand out from the rest. Statistical testing of the results indicates the same, i.e. the separation of HKE from other varieties. Chi-

square tests[91] on raw frequencies of the progressive and the actual number of words in a corpus (M-coefficient) indicate statistically significant differences between IrE and IndE, and SinE and HKE ($p < 0.001$ for both), i.e. the varieties with most and least frequent use of the progressive. There are no statistically significant differences between the other six varieties. However, when using the raw frequencies of the progressive and the number of verb phrases in a corpus (V-coefficient), the indications of the statistical testing are slightly different. The significance of the difference between IrE and IndE is diminished to $p = 0.05124$, but on the other hand, a significant difference arises between JamE and PhiE ($p = 0.04026$).

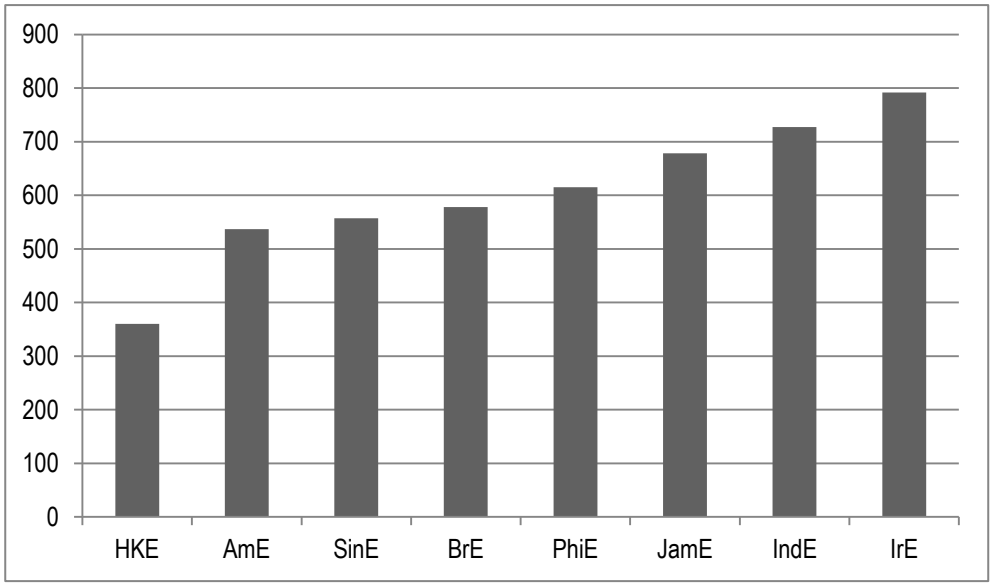

Figure 7. Frequency of the progressive in World Englishes, V-coefficient.

Further chi-square tests on the data presented in Figure 7 indicate that AmE, SinE, BrE and PhiE form a cluster with very similar frequencies. SinE is reported to follow the major standard varieties (more precisely, BrE) closely (Sharma 2009: 182-8; see also subsection 5.3 below). Most other varieties thus represent overuse of the progressive as compared to the two major varieties, whereas in HKE, the

[91] See Appendix E for the raw frequencies and the results of the chi-square tests performed on the data.

progressive is clearly underused. It is noteworthy, however, that the variety with most frequent use of the progressives is not an OC variety but an IC variety, IrE. Furthermore, there is no grouping of IC and OC varieties separate from each other, but rather an even dispersion of varieties across the scale of frequency of the progressive.[92] The following section discusses the variation in the use of the progressive in the eight varieties in more detail, and presents possible reasons for the use of the progressive in a number of the varieties included in the present study.

5.3 The role of substrate transfer

There are a number of possible reasons that may explain the differences in the frequency of the progressive in World Englishes. The most frequently discussed and generally accepted reason lies in substrate transfer from local languages. Other possibilities discussed in previous literature are universals of second language acquisition, or so called 'angloversals' (see e.g. Sharma 2009, Davydova 2012). Here, only one of the proposed reasons is discussed in detail: substrate transfer.[93]

Filppula (2003, 2008) reports that the use of the progressive in IrE is relatively free, and proposes substrate transfer from Irish, as well as other Celtic languages and dialects of Celtic Englishes, as the main source of influence. According to Filppula (2003), the Celtic influence on the use of progressive in English (in general) is supported by four criteria: (i) geographically, the Celtic progressive form is the one closest to the English progressive form, (ii) chronologically, the Celtic progressive form precedes the English progressive form, (iii) socio-historically, the Celtic-English contact favoured language shift from the Celtic substrate to English, and (iv) in general, Celtic influence on other features of English grammar has been shown to have taken place. Celtic-influenced non-standard uses of the progressive are found, in addition to IrE, in Hebridean English, Welsh English, and Scots and Scottish English, as well as in some dialects of English spoken in the western and northwestern parts of the British Isles (formerly Celtic-speaking areas). Filppula

[92] Collins (2008) and Hundt and Vogel (2011) report similar findings without grouping of IC or OC varieties as regards the frequency of the progressive.

[93] Possible reasons affecting the frequency of the progressive in World Englishes are discussed from a wider perspective in Meriläinen et al. (fc.).

and Klemola (2012: 1689-90) provide a list of non-standard uses of the progressive that are typical of IrE and other Celtic-influenced Englishes:

a) the progressive may be used with stative verbs

(5.1) There was a lot of fairies long ago […] but I*'m thinkin'* that most of 'em are vanished. (cited in Filppula 1999: 89)

b) the progressive may be used with dynamic verbs to express habitual activities or states of affairs (BrE prefers the simple forms)

(5.2) …but there, there's no bogland here now.

[Interviewer: Yeah. = And do people go up there to cut turf?]

They *were going* there long ago but the roads got like, like everything else, they got a bit too-o rich and […]. (cited in Filppula 2003: 162)

c) the progressive may be used with the auxiliaries *would/'d/used [to]* to indicate habitual activity (BrE prefers the simple forms)

(5.3) So, when the young lads*'d be going* to bathing, like, they'd have to go by his house, and they used to all, he u', he loved children. (cited in Filppula 2003: 163)

d) the progressive may be used with the auxiliaries *do/does* and *will/'ll*.[94]

(5.4) Yeah, that's, that's the camp. Military camp they call it […] They *do be shooting* there couple of times a week or so. (cited in Filppula & Klemola 2012: 1690)

(5.5) [T]his fellow now, Jack Lynch, that's going to come into power now, that he'll, he*'ll be forgetting* the North. (cited in Filppula & Klemola 2012: 1690)

Filppula (2003), having provided a similar list of IrE uses of the progressive, remarks that not all of the uses listed are unique to IrE (e.g. extended use of the progressive with stative verbs has been reported in a number varieties throughout the world, see e.g. Kortmann and Lunkenheimer 2013), and on the other hand, that not all of the uses listed have an exact parallel in Irish (e.g. progressive form of

[94] The use of the progressive with *do/does* is restricted to IrE, and the construction is used to express habitual events or states. The use of the progressive with *will/'ll* is common in other varieties, as well (cf. section 6.3 in the present study).

BELIEVE or KNOW). Language contact situations often produce features which both reproduce and overgeneralise syntactic patterns from the substrate language (Filppula 2003: 164). What kind of functions the progressive is used in present-day IrE is discussed in Chapter 8. What the present data does attest is a very frequent use of the progressive in IrE (V=792).

Another variety with frequent use of the progressive is IndE (V=727). Extended use of the progressive is one of the most prominent features related to IndE in previous studies (e.g. Rogers 2002, Balasubramanian 2009, Kortmann and Lunkenheimer 2013). However, Davydova (2012: 371, 374) points out that the extended use of the progressive in IndE is restricted to the acrolectal level, while on the basilectal level progressives are rare. Furthermore, the situation is even more complicated when the fact that speakers with different background languages may use the progressive differently is taken into account. Nevertheless, both Davydova (2012) and Sharma (2009) propose that the extended use of the progressive in IndE can be traced back to the Hindi substratum, the primary substrate language of Northern India.[95] In Hindi, an overt auxiliary *rahna* ('remain') is required to mark the progressive aspect, while *–ta* marks the non-progressive and habitual aspects (Sharma 2009: 174). What is crucial here is the fact that these aspectual markers are obligatory, i.e. Hindi requires either *rahna* or *–ta* to co-occur with verbs to denote the desired aspectual values. This leads to speakers of IndE experiencing a need to mark both perfective and imperfective aspect with overt markers, and as *–ing* is the only available imperfective marker in StE, speakers of IndE reinterpret it a global marker for imperfectivity (ibid.: 185).

In addition to IndE, Sharma (2009) discusses the use of the progressive in SinE. Although the background for the use of the progressive in SinE is rather similar to that found in IndE, there is a fundamental difference that leads to the great difference in the frequency of the progressive in the two varieties. Similarly to Hindi, the substrate languages of SinE[96] require an overt progressive marker, but the difference is that non-progressives are marked optionally and in rather restricted environments, and habituals are not marked at all (see Table 20 below). Thus due to the fact that imperfective marking is, at least in part, optional in the

[95] Davydova (2012: 371) mentions both Hindi and Punjabi as the main substrate languages of her data, while Sharma (2009) draws her data from Hindi speakers, some of them having Gujarati or Punjabi as additional languages. The present data has a more varied substrata, the majority of speakers having Kannada or Marathi as mother tongue. Whether the present data with its non-Hindi substrate languages also shows extended use of the progressive is discussed in Chapter 8.

[96] Singapore Mandarin, Cantonese, Teochew, Hokkien and Malay. See section 3.3.

Table 20. Aspect marking in Singaporean substrate languages (Sharma 2009: 177).

Aspect	Singapore Mandarin	Cantonese	Teochew	Hokkien	Malay
Progressive	*zai*	*gan*	*do*	*tja*	*sedang*
Non-progressive	*(-zhe)*	*(zyu)*	*(do)*	-	-
Habitual	-	-	-	-	-

substrate languages of SinE, there is no pressure on the speakers of SinE to extend the use of *–ing* to other imperfective situations. As Sharma (2009: 182-8) points out, and as is shown in the present data, SinE use of the progressive is rather similar to StE usage.[97]

Ansaldo (2004: 137-8) combines the idea of language contact with that of universality by suggesting that the progressive aspect in SinE is the result of "contact-induced change that reflects shift based on substratal influence and 'universal' innovation process typical of contact situations", i.e. a hybridization. Thus, in SinE, we find morphological reduction due to the isolating typology of the substrate (substratal influence) and reanalysis of temporal adverbials into aspect markers (universal innovation process). In more concrete terms, SinE shows loss of *–ing*, optionality of the auxiliary BE, and the temporal adverbial *still* has been given aspectual force. Example (5.6) below shows a rather stripped way of expressing the progressive aspect in SinE:

(5.6) Still rain now. 'It's raining.' (Ansaldo 2004: 137)

However, due to the focus of the present study, auxiliary BE + present participle, whether 'progressives' such as in (5.6) occur in the present data is not investigated here. Use of the auxiliary BE is discussed in sections 6.5 and 6.6.2, and the co-occurrence of temporal adverbials with the progressive in subsection 6.6.3. Nevertheless, it should be pointed out that it is possible that ICE-SIN contains

[97] Frequency of progressives is very similar in BrE and SinE (see Figure 7 above); functions of the progressive also show similarities between BrE and SinE (see Table 53 in section 8.6).

more cases of the progressive aspect than merely those represented by the progressive form proper.

Finally, perhaps the most intriguing question concerns the frequency of the progressive in HKE – why is the progressive so infrequent in HKE (V=360)? This question is difficult, if not impossible, to answer within the scope of the present study, mainly due to the lack of previous studies on the progressive in HKE. As pointed out in the preceding subsection, the results of Collins (2008) are not comparable to this study due to a fundamental difference in the retrieval process, whereas Hundt and Vogel (2011) do not include HKE in their selection of varieties, and while Hundt (2009a) does include ICE-HK in her data, she focuses on the progressive passive instead of the full paradigm. However, Meriläinen et al. (fc.) include data from ICE-HK[98] and ICLE-HK: results from the learner data confirm that the progressive is very infrequent in HKE. Other than these studies, I have not been able to locate studies on the progressive in HKE. However, as it is known that the speakers included in ICE-HK have Cantonese as their first language (Bolt and Bolton 1996: 199), the underlying substratal situation could be depicted as similar to that in SinE, i.e. the progressive aspect is marked by *gan*, while other imperfective situations are overtly marked only in restricted environments (see Table 20 above). According to Yip and Matthews (2000: 99), the progressive marker in Cantonese in optional, which could perhaps cause the speakers of HKE to experience no need to mark progressivity overtly in English, either. In fact, the present data includes a fair amount (approx. M=20) of instances of the auxiliary BE combined with the uninflected form of the main verb, consider (5.7) and (5.8):

(5.7) we *are spend* most time on listening rather than speaking (ICE-HK, S1A-077)

(5.8) Oh what *are* you *study*? (ICE-HK, S1A-005)

Whether this 'corrupt' form truly is a progressive is unclear, as another plausible explanation, offered by Hung (2012: 128), is that the auxiliary BE is in fact used for emphasis (instead of StE DO). However, both (5.7) and (5.8) above seem to fit a progressive reading better than an emphatic one, especially the interrogative in (5.8). Hung (ibid.) also points out that the auxiliaries BE and DO are not used in Chinese, which offers yet another possible explanation for (5.8): the speaker may

[98] Meriläinen et al. (fc.) is partly based on the present data, including parts of ICE-HK, and thus should not be compared to the present results.

have mixed the auxiliaries and uses BE instead of the more appropriate DO. Nevertheless, that these instances are in fact progressives is further backed up by the fact that also the past participle is found uninflected in the ICE-HK data, see (5.9):

(5.9) Yes uh also uh two other groups are hav *have approach* me (ICE-HK, S1A-005)

Hundt (2009a: 293) also finds the progressive passive combining BE *being* with an uninflected past participle in mainly spoken HKE, as well as SinE and IndE. Yet another explanation is an error in the transcription, ie. the ending *–ing* may have been lost in the transcribing process. In the present study, the 'corrupted' progressive forms are not included in the calculation of the frequency of the progressive, as the starting point of the study was to search for the forms of BE combined with the *–ing* form of the main verb.[99] Nevertheless, the low amount of BE + V*ing* and the instances of BE + V in ICE-HK suggest that the expression of the progressive aspect may be more complex in HKE than seems at first: further research is necessary to cover all possible progressives in HKE.

The present section discussed the role of substrate languages on the frequency of the progressive in four World English varieties. Two of the varieties discussed showed high frequency of the progressive (IrE and IndE), one very low frequency (HKE), and another had a frequency very close to BrE and AmE (SinE). In each case, the use of the progressive aspect in the substrate languages seemed to have an effect on the local English variety. In IrE, the progressive is used rather freely in contexts where StE prefers the simple forms. In IndE and SinE, a mis-match between the aspect systems of the substrates and the superstrate led to different outcomes: speakers of IndE interpret the English *–ing* ending as a global marker for the imperfective aspect, thus using it frequently, while speakers of SinE do not feel a need to extend the use of *–ing* and thus approximate their use of the progressive to the StE use. A similar mis-match and optionality of progressive marking has possibly led to underuse of the progressive in HKE, although it remains unclear why exactly this has happened.

Summarising the frequency of the progressive in World Englishes, IrE and IndE have the highest frequencies, while the progressive is clearly underused in HKE. According to the present data, OC Englishes do not overuse the

[99] The possible progressive forms with the uninflected present participle were found during a closer examination of the text files, i.e. during the process of tagging all verb forms in the data (cf. section 4.1.2).

progressive, although BrE and AmE do occur towards the lower end of the frequency. Overall, there are very little differences between the varieties investigated here, excluding HKE and IrE, which are significantly different. The present section discussed substrate transfer as a possible reason for the distribution of progressives in the present data.

6 Morphosyntactic variation

In this chapter, the progressive is studied from a number of different viewpoints, all related to the grammatical environment that the progressive occurs in. The selection of features is largely based on previous studies, in particular Smitterberg (2005) and Collins (2008), who mostly base their selection on Smith (2002). Following these studies, the focus of the present section is on morphosyntactic variation of the progressive within the verb phrase, i.e. tense (present vs. past), the perfect progressive (*have/has/had been* + V*ing*), modal auxiliaries (modal auxiliaries + V*ing*), and voice (*being* V*ed*). As Smitterberg (2005: 115) states, these features may have influenced the frequency of the progressive in the past centuries: one channel for the spread of the progressive is, quite naturally, the extension of the construction to new forms. Mair (2006: 88-89) discusses the same channel of spread, and proposes that one of the routes of spread is the creation of new forms (see section 5.1 above). Also Kranich (2010: 113, 121-8) discusses the grammaticalization of the progressive, providing evidence from the period of Modern English. Based on previous studies and her own data, Kranich establishes that the spread of the progressive has not been influenced by the extension of the progressive paradigm, but rather by the increasing use of the already established forms. Smith (2002: 318-9) locates the spread of the progressive to the category of present progressives: the proportion of present progressives increased by 30% between the 1960s and the 1990s, while the proportion of past progressives decreased by approximately 10%.

One of the aims of the present chapter is to see whether there is evidence for the progressive paradigm having expanded to new forms in any of the varieties studied. Table 21 below presents the overall distribution of the progressive across different forms (simple/modal aux./perfect/passive/fragmentary). It is clear that the progressive occurs most commonly in the simple contexts (*am/are/is/was/were* and their contracted forms): 84% of all progressives in the data appear in sentences such as (6.1) and (6.2) below:

(6.1) Wait how come you *'re asking* me those questions (ICE-PHI, S1A-065)

(6.2) Yeah I *was talking* to him last Saturday night (ICE-IRL, S1A-067)

Table 21. Distribution of progressives across different forms in World Englishes.

Form	World Englishes
Simple progressives	84.0% (5,524)
Modal progressives	5.1% (325)
Perfect progressives	3.3% (208)
Passive progressives	1.1% (70)
Fragmentary progressives	3.4% (218)
TOTAL	100% (6,345)

The present results conform to those obtained in previous studies discussed above: the progressive is not frequent in the more complex combinations. Modal progressives is the largest group of complex progressives in World Englishes, with a proportion of approximately 5%, while perfect progressives are found in 3%, and passive progressives in as little as 1% of all progressives. These results indicate that the spread of the progressive into World Englishes has not taken place through extension of the progressive paradigm, at least in spoken language. Indeed, the complex progressives may be more frequent in written registers (see e.g. Collins (2008: 233-4) or Smitterberg (2005: 134)).

In the following, section 6.1 discusses the use of present and past tense forms in the simple progressive, while sections 6.2 to 6.4 focus on the different complex combinations in which the progressive occurs. The information from the separate subsections is combined in a section dealing with all the possible verb phrase patterns that the progressive occurs in (section 6.5). Other features investigated in the present chapter are clause type, contraction, and temporal specification of the progressive (section 6.6). These features have been shown to play an important role in the study of the progressive, as they have contributed to the spread of the progressive.

In addition to morphosyntactic variation within the progressive paradigm, the data used in the present study enables the investigation of variation between the progressive and the non-progressive (see section 4.4). However, due to the amount of manual work necessary, the investigation of the progressive vs. non-progressive paradigm is quantitative in nature, and does not attempt any detailed analyses on the non-progressive.

6.1 Tense

Due to the high proportion of simple progressives, this section focuses on tense choices of simple progressives only: the low proportion of complex and fragmentary progressives entails that the tense choices made in those categories do not affect the overall distribution of present and past tense progressives.[100] Progressives in combination with modal auxiliaries, the perfect, or the passive voice are discussed below: reference to tense choices are made where relevant. The number of progressives under scrutiny in this section is 5,524, i.e. 84% of all progressives found in the data, and examples were provided in (6.1) and (6.2) above.

Previous studies provide information mostly on BrE: Smitterberg's (2005: 118) data on 19th-century written BrE shows that a present-tense progressive form is found only slightly more often than a past-tense progressive (55% : 45%), but when individual genres are examined, it seems that some of them actually prefer the past tense (e.g. Fiction). Kranich (2010: 126-7) shows that the proportion of present tense forms has slightly increased towards the end of the millennium: it is only during the 19th century that the proportion of present tense progressives surpasses that of past tense progressives. Similarly, Smith's (2002: 318-9) results on late 20th-century written BrE report a rather steep increase in the proportion of present tense progressive at the cost of past tense progressives. Römer (2005: 62) provides information on spoken present-day BrE, showing that as much as 68% of the progressives in her data are in the present tense. Collins' (2008) data draws from similar sources as this study, but incorporates both written and spoken language. Unfortunately, Collins does not provide data on the use of the progressive in speech in the individual varieties, but a table comparing speech and writing in general (Table 2d, 2008: 234) shows that the progressive occurs in the present tense more often in speech than in writing (c. 63% speech - c. 35% writing).[101]

Figure 8 below portrays the proportion of present tense progressives in the individual varieties. The present tense is favoured over the past tense in all

[100] In an earlier manuscript of this study it was shown that this truly is the case: the overall distribution of present and past tense progressives is highly dependent on simple progressives.

[101] The difference between speech and writing may arise from the fact that referring to what is happening *now* is more common in speech and requires the present progressive (see e.g Kranich 2010: 127).

varieties: on average, 70% of the progressives studied were in the present tense. In relation to results obtained in previous studies - most of which focus on written BrE - it seems that the present tense is very strongly favoured in World Englishes on average. As Figure 8 shows, IC Englishes distribute present and past tense forms clearly more evenly than do OC varieties: the difference is statistically significant ($\chi^2(1) = 217.3522$, $p < 0.001$). Contrasting the proportion of present tense progressives to the proportion of present tense non-progressives[102] reveals a statistically significant difference overall ($\chi^2(1) = 20.076$, $p < 0.001$): the present tense is more frequent with the non-progressive (see Table 22 below).

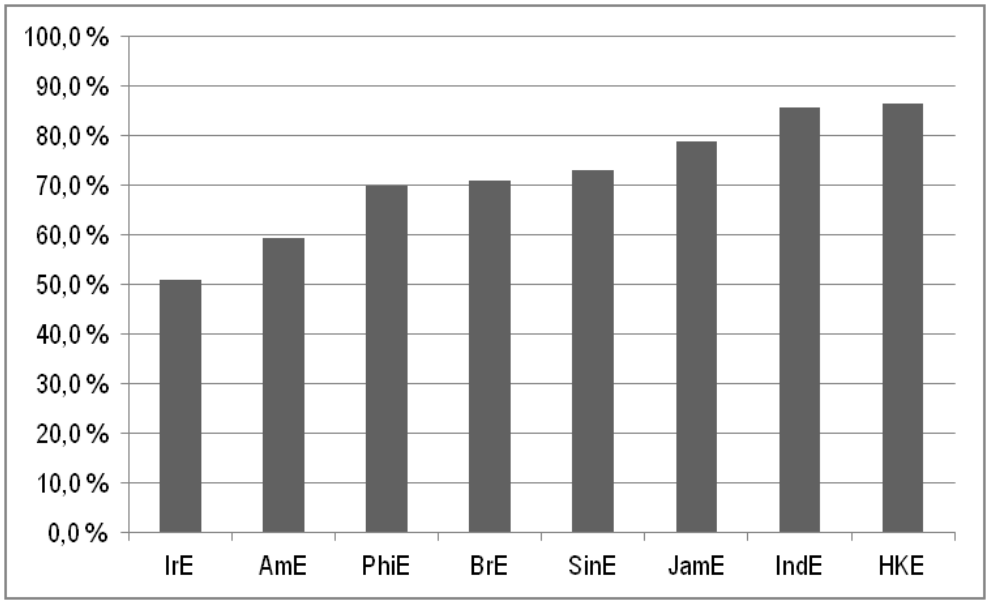

Figure 8. Proportion of present tense simple progressives in World Englishes.

Looking at the individual varieties, a more complex picture arises. In IrE and AmE, the difference between the proportions of present and past tense use is less significant (51% of all progressives are in the present tense in IrE, and 60% in AmE). In contrast, HKE, IndE, and JamE, in particular, strongly opt for the present tense with the progressive form (more than 75% of all progressives are in

[102] Here, the statistical testing was performed on present tense simple and passive forms, since during the data retrieval and tagging of verb phrases no attention was paid to non-progressive passive forms, i.e. they were not coded separately. Because of this, the progressive vs. non-progressive paradigm is not discussed in relation to passive forms in section 6.4 either.

the present tense). BrE, SinE, and PhiE are situated in the middle of the two ends, with approximately 70% of all progressives occurring in the present tense. Comparison of the IC and OC varieties brings out a clear difference between the two groups - in the IC varieties, circa 60% of all progressives are in the present tense, while the corresponding figure for the OC varieties is as high as 79%. The high frequency of present tense progressives in OC Englishes suggests that the use of progressives may be more restricted to its most prototypical environment, referring to situations in the present moment, when the use of the progressive is almost obligatory.[103] On the other hand, the difference may arise from the data: as the amount of data included in the present study is limited, the topics of the conversations included may affect the occurrence of certain linguistic features.[104]

The greatest amount of variation in the use of the tense forms is found in IrE and AmE, while IndE and HKE strongly favour present progressives. This could be considered to indicate that the progressive is more firmly integrated into IrE and AmE than IndE and HKE.[105] In other words, the speakers of IrE and AmE are able to use the full potential offered by the two tenses, whereas in IndE and HKE the progressive is used in a more restricted manner. In order to verify whether the present tense forms in IndE and HKE (and possibly also JamE) are truly as frequent as they appear to be with the progressive, we need to investigate how present tense is distributed within non-progressives. Interestingly, the two varieties with statistically highly significant differences in the proportion of present tense forms are IrE and IndE (see Table 22 below). Furthermore, the data shows that the progressive in IrE underuses present tense forms, while IndE overuses them. Thus the argument concerning the use of the two tenses in, on the one hand, IrE and AmE and, on the other, IndE and HKE is, at least partially, confirmed: speakers of IrE use the progressive more flexibly in both tenses, while speakers of IndE are more confined to the present tense forms. However, the distribution of

[103] The functions of the progressive in World Englishes are discussed in Chapter 8.

[104] For instance, the progressive may be more frequent in narratives and storytelling than in discussions on cooking or one's habitual activities.

[105] Throughout the present chapter, the notion of integration (see Smitterberg 2005: 57-58) is used to discuss the differences between the individual varieties as regards morphosyntactic variation. Smitterberg (2005) investigated the diachronic integration of the progressive into Modern English, whereas the present study applies Smitterberg's study to evaluate the different levels of integration in individual World Englishes. In the present study, the level of integration is related to the frequencies of a given feature in individual varieties. Integration of the progressive to World Englishes is further discussed in section 6.7.

present tense forms in AmE and HKE does not differ as regards the progressive and the non-progressive.

Table 22. Progressive vs. non-progressive, present tense.[106]

Variety	Present progressives[107]	Present non-progressives	Diff. in % units
BrE	70.8% (550)	73.6% (8,321)	-2.9
IrE	51.2% (499)	59.6% (6,312)	-8.4 ***
AmE	59.0% (406)	62.2% (7,372)	-3.2
JamE	79.3% (570)	78.3% (8,037)	+1.0
IndE	85.6% (642)	80.1% (7,419)	+5.5 ***
PhiE	70.5% (510)	69.9% (7,517)	+0.7
SinE	73.1% (469)	76.7% (8,118)	-3.6 *
HKE	86.3% (276)	88.2% (7,367)	-2.0
TOTAL	70.1% (3,922)	72.9% (60,480)	-2.8 ***

A possible explanation for the underuse of the present progressive in IrE may lie in the functions of the progressive: IrE uses communication verbs rather frequently and mostly they occur in the past tense, as in (6.3) below:

(6.3) And she *was saying* well I can make my voice authoritive if I want to (ICE-IRL, S1A-028)

Communication verbs are further discussed in section 7.2 below.

[106] Asterisks show the level of statistical significance: p = 0.05 is indicated by *, p = 0.01 is indicated by **, and p = 0.001 is indicated by ***. See section 4.5 for more information.

[107] The figures in the table are based on the number of present tense simple progressives and progressive passives.

To sum up, the present tense is clearly favoured over the past tense in progressive verb phrases. This preference prevails in all varieties studied, although there is considerable variation from one variety to another. The high proportion of present tense progressives conforms to results obtained in previous studies, and may also be regarded as evidence for the increasing trend of present tense progressives in the late 20th century, observed by Smith (2002) and Kranich (2010). Moreover, the fact that present tense progressives are more frequent in OC varieties may be regarded as indication that the progressive is not as well established in OC varieties as it is in IC varieties.

6.2 The progressive with modal auxiliaries

The combination of a modal auxiliary and the progressive is reported to be quite rare and, according to Smitterberg (2005: 134), its use decreased significantly during the 19th century. Smitterberg offers several possible reasons for the declining trend, including the decrease of the use of modal auxiliaries in general. Smitterberg also suggests that the progressive may, in some contexts, express modal values in itself, thus the need to apply modal auxiliaries diminishes (2005: 134). Furthermore, it is possible that some combinations of a modal auxiliary and the progressive may not have been accepted – Smitterberg (2005: 137) reports on constraints on what meanings could be expressed by modal progressives in the 19th century. It is possible that such constraints, or new ones, are in effect in late 20th-century English, as well. On the other hand, Smith (2002) and Kranich (2010) report on an increase in the use of the progressive with modal auxiliaries in BrE during the 20th century. Collins and Yao (2012) investigate the use of modals and quasi-modals in thirteen varieties of English[108], and report that AmE seems to be leading the way in the increase of the quasi-modals and the decrease of the modals investigated. From this perspective, it is interesting to see how modal auxiliaries are used with the progressive in World Englishes in the late 20th century. Sentences (6.4) and (6.5) exemplify the use of modal progressives:

(6.4) You *'ll be starting* at eight o'clock (ICE-PHI, S1A-053)

(6.5) So she *must have been going* to Utah when she was like fifteen (SBCSAE, 0036)

[108] The modals investigated by Collins and Yao (2012) are *must, should, will* and *shall*, and the quasi-modals *have to, have got to, be going to* and *want to*.

Figure 9 below shows that the use of modal auxiliaries with the progressive is indeed rather low: they account for approximately 5% of all progressives in the data.[109] The one variety very different from all the others is IndE in which the progressive occurs with modal auxiliaries in as much as 12% of all progressives. The amount of progressives with modal auxiliaries in IndE is staggeringly high as the variety with second most frequent use is PhiE with a proportion of less than 6%. Least frequent use is found in BrE and AmE, and on the other hand, in HKE (less than 3% in each). The distribution of progressives with modal auxiliaries in IC and OC varieties is statistically non-significant when IndE is excluded ($\chi^2(1) = 1.4853$, $p = 0.2229$), thus IndE is truly set apart from the other varieties in this regard.

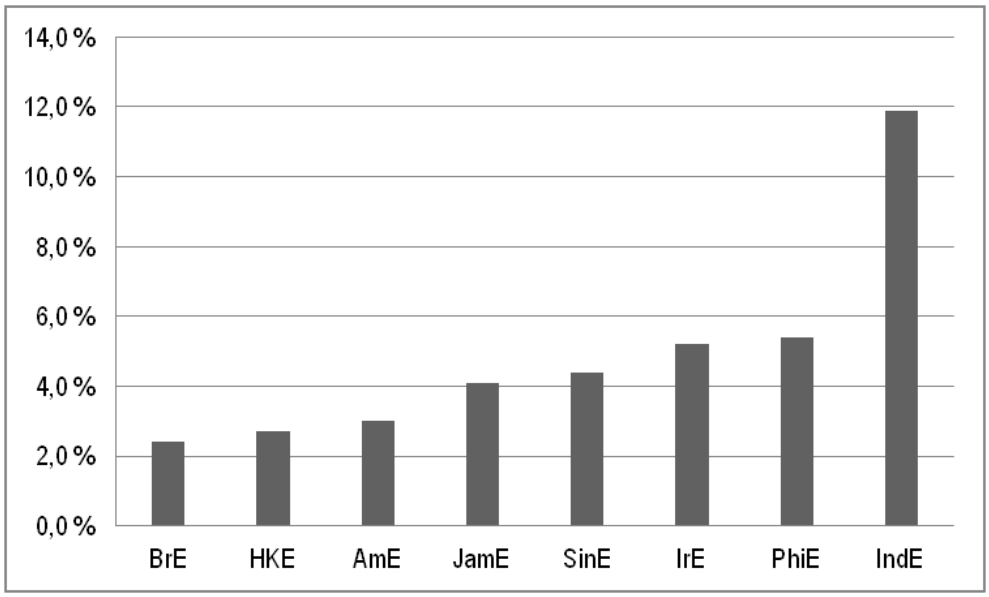

Figure 9. Distribution of the progressive with modal auxiliaries in World Englishes.

The low frequency of the progressive with modal auxiliaries in the IC varieties, especially BrE and AmE, may be seen as evidence of the decreasing use reported by Smitterberg (2005: 134). Furthermore, the high frequency found in IndE could be the result of colonial lag – "the retention of earlier linguistic features" (Marckwardt 1958, cited in Hundt 2009b: 13). In other words, colonial lag refers to

[109] Raw frequencies are presented in Table 24 in the present section.

a feature that in the receiving variety is used as it was used in the donor variety at the time of contact.[110] Kranich (2010: 126) provides quantitative data on the use of the progressive with modal auxiliaries in the Modern English period, i.e. the time when English was transported to India. Unfortunately, the raw frequencies and the percentages are very low throughout the period covered, ranging from 1% to 9%. Moreover, the proportion of progressives with modal auxiliaries in Kranich's data fluctuates from a half-century to another, so that for the latter part of the 17th century, the proportion of progressives with modal auxiliaries is 9%, whereas in the following period, the early 18th century, the corresponding figure is 3%. Thus, the attempt to try to draw any conclusions regarding the colonial lag theory on the basis of Kranich's data is futile. Naturally, it is also possible that it is the more recent varieties where the change is taking place – IndE and PhiE could be seen as innovators on their way to more frequent use of modal progressives. It is impossible, however, to come to any definite conclusions in the scope of the present study, as a diachronic set of data would be needed to sketch out the development taking place in the past decades.

Similarly to the discussion on tense in the previous section, the present data enables the comparison of the use of modal auxiliaries with the progressive as well as the non-progressive. Overall, the use of modal auxiliaries is less frequent with progressive than non-progressive verb forms, which suggests that the combination of modal auxiliaries and the progressive may not be thoroughly accepted in all varieties under investigation (see Table 23 below). Approximately 12% of non-progressive verb phrases include a modal auxiliary, while the corresponding figure for progressive verb phrases is 5%. Thus Smitterberg's (2005: 134) note about the combination of a modal auxiliary and a progressive form seems to be true - they are rather rarely found together. BrE, especially, exemplifies this matter well. Although the amount of modal auxiliaries combined with non-progressive verbs is among the highest found in the present data (c. 13%), the amount of modal auxiliaries with the progressive form is very low, in fact, the lowest of all eight varieties (c. 2%). HKE follows closely to BrE, while AmE has a more moderate proportion of non-progressives with a modal auxiliary. PhiE and especially IndE, on the other hand, show high frequencies for the use of modal auxiliaries combined with the progressive form. In IndE, the amount of modal auxiliaries is high with non-progressive verb forms (c. 14%), as well, while in PhiE non-

[110] For a thorough discussion on colonial lag, see Hundt 2009b.

progressive verb forms combine with modal auxiliaries in significantly fewer instances (c. 10%). Thus it could be argued that the amount of modal progressives is even more noteworthy in PhiE than seems at first.

Table 23. Modal auxiliaries with progressive and non-progressive verbs.

	Modal progressives	Modal non-progressives	Difference in % units
BrE	2.4% (20)	13.3% (1,821)	-10.9 ***
IrE	5.2% (56)	12.1% (1,504)	-6.9 ***
AmE	3.0% (23)	10.4% (1,413)	-7.4 ***
JamE	4.1% (35)	10.7% (1,268)	-6.6 ***
IndE	11.9% (106)	14.1% (1,601)	-2.2
PhiE	5.4% (43)	9.7% (1,192)	-4.3 ***
SinE	4.4% (32)	12.7% (1,579)	-8.3 ***
HKE	2.7% (10)	12.7% (1,277)	-10.0 ***
TOTAL	5.1% (325)	11.9% (11,655)	-6.8 ***

The exceptionality of PhiE is strengthened when modality in general is taken into account, by combining the proportion of modal auxiliaries in progressive and non-progressive verb phrases in each variety.[111] For IndE, the proportion of verb phrases with a modal auxiliary in general is high at approximately 14% of all verb phrases, and thus the high frequency of modal auxiliaries in progressive verb phrases is not as surprising as it is with the same features in PhiE: the 6%-proportion of progressives with modal auxiliaries is contrasted to a 9.5%-

[111] In the present data, the proportion of modal auxiliaries in progressive and non-progressive verb phrases combined is as follows: PhiE & AmE 10%, JamE 11%, IrE 12%, SinE, HKE & BrE 13%, IndE 14%.

proportion of modality in general. Overall, the proportion of verb phrases with modal auxiliaries in World Englishes is approximately 11.5%.

The high frequency of modals, especially in progressive verb phrases, in IndE is somewhat surprising, as previous studies do not report on any major differences between the use of modals in IndE and BrE. A study by Katikar (1984, cited in Sailaja 2009) on the use of modals in written IndE indicates "very marginal" differences between IndE and BrE. Similarly, Balasubramanian (2009: 110-1) reports that the use of modals in IndE is very similar to BrE and AmE. In their spoken data, Collins and Yao (2012: 45, Table 5), on the other hand, find a significant difference in the use of modals between BrE and IndE, but also between SinE and IndE, SinE being the variety with the highest frequency of modals. On the basis of the present study, however, it seems that in spoken IndE there definitely are differences regarding the use of modal auxiliaries, especially with progressives, compared to all other varieties studied. Looking more closely at the distribution of the different modal auxiliaries co-occurring with the progressive in IndE, we find that a number of the modals stand out: *may, must* and *will*, perhaps also *shall*, are clearly more frequent in IndE than in any other variety. For instance, there are 15 instances of *must be *ing* in IndE, while the corresponding figure for all other varieties combined is only four. A larger set of data and a more detailed investigation of the modals would most likely offer explanations to these findings. Unfortunately, it is not possible to discuss this interesting matter in more detail within the scope of the present study, but there are clear indications that further research into the use of modal auxiliaries in World Englishes is needed.

It is now time to turn the attention which modal auxiliaries co-occur with the progressive in the present data: the auxiliaries investigated are CAN – *can* (*have*)/*could* (*have*), MAY – *may* (*have*)/*might* (*have*), MUST – *must* (*have*), SHALL – *shall* (*have*)/*should* (*have*), and WILL – *will* (*have*)/*would* (*have*)). It must be noted that not all instances of *could*, for instance, are the true past form of *can*, but used as conditional forms thus possibly referring to a future situation. Nonetheless, they are discussed together as their meanings are usually close to each other. Table 24 below shows that, overall, WILL is the most frequent modal auxiliary co-occurring with the progressive, occurring in as much as 73% of all progressive verb phrases with a modal auxiliary. The proportions of MAY, MUST and SHALL are more evenly distributed at approximately 8%, while CAN is found in only 4% of all progressives with a modal auxiliary. The frequency of *can* and *could* seems to be the most remarkable difference when the present results on the combination of modal auxiliaries and the progressive are compared to those provided by Biber et al. (1999: 486) on the

frequency of modal auxiliaries overall: *can* and *could* co-occur with the progressive less frequently than with the non-progressive. On the other hand, *must* occurs commonly with the progressive.

Looking at the distribution of the five modal auxiliaries in the individual varieties (Table 24 below), we find that the use of modal auxiliaries varies to a great extent from one variety to another. All five modal auxiliaries are found in three varieties only (IrE, IndE, SinE), whereas the other varieties do not make use of all five auxiliaries. For instance, in JamE the distribution of modal auxiliaries with the progressive is rather even, except for MUST which was not found in ICE-JA in combination with the progressive. On the other hand, PhiE and HKE show the least amount of variation - close to 90% of all modal auxiliaries are either *will* or *would*. HKE, especially, almost exclusively uses *will* (9 out of 10 occurrences). A possible explanation for the result in HKE in particular, but other varieties also, may be the low raw frequencies – there might not be enough data to bring out the true distribution of modal auxiliaries co-occurring with the progressive.

When similar data concerning the use of modal auxiliaries is drawn from the non-progressive verb phrases, it becomes evident that the distribution of modal auxiliaries is more stable in the non-progressive than the progressive. Approximately 80% of modal auxiliaries used with non-progressive verbs are either WILL or CAN, while the final 20% are distributed between MAY, MUST and SHALL. All five modal auxiliaries are used in all eight varieties. The distribution of modal auxiliaries is surprisingly even among the eight varieties – the only discernible difference seems to be that the OC varieties use SHALL more than the IC varieties. Mair (2006: 101) reports on a decrease in the use of *must* and *shall* in BrE and AmE; at least for *shall*, this seems to be true based on evidence from the ICE corpora. Collins and Yao's (2012: 43-45) results show that, overall, modals are more frequently used in OC varieties, suggesting that the IC varieties are in the lead of the decreasing trend. Furthermore, Collins and Yao (ibid.) are able to show that AmE is the most advanced variety with regard to the decreasing use of modals and the increasing use of quasi-modals. As regards the present data, AmE is characterized by low use of modal auxiliaries in both paradigms investigated. Overall, it could be argued that the use of modal auxiliaries with non-progressive verbs is stable, while using them with progressive verbs does not portray any strong recurring patterns, which may indicate that the combination of the progressive and modal auxiliaries has not been fully established in World Englishes. Another possible reason is the overall low frequency of modal auxiliaries combined

Table 24. Distribution of modal auxiliaries combined with the progressive in World Englishes.

Variety	CAN	MAY	MUST	SHALL	WILL	Total
BrE	0	25.0% (5)	5.0% (1)	15.0% (3)	55.0% (11)	100% (20)
IrE	5.4% (3)	8.9% (5)	5.4% (3)	5.4% (3)	75.0% (42)	100% (56)
AmE	0	0	17.4% (4)	8.7% (2)	73.9% (17)	100% (23)
JamE	17.1% (6)	11.4% (4)	0	17.1% (6)	54.3% (19)	100% (35)
IndE	0.9% (1)	8.5% (9)	14.1% (15)	5.7% (6)	70.8% (75)	100% (106)
PhiE	0	4.7% (2)	0	7.0% (3)	88.3% (38)	100% (43)
SinE	6.3% (2)	6.3% (2)	6.3% (2)	3.1% (1)	78.1% (25)	100% (32)
HKE	0	0	0	10.0% (1)	90.0% (9)	100% (10)
Overall	3.7% (12)	8.3% (27)	7.7% (25)	7.7% (25)	72.6% (236)	100% (325)

with the progressive aspect in the present data. A larger dataset might bring out a more stable distribution also within the progressive paradigm.

In the following, the distribution and frequency of the five modal auxiliaries are discussed in more detail. However, the meanings conveyed by the modals are not included in the discussion, unless it is relevant.

CAN - *can* (*have*)/*could* (*have*)

The amount of *can* and *could* co-occurring with the progressive is very low in the present data: in four of the varieties, CAN was not found at all (RF=12; see Table 24 above). Sentences (6.6) and (6.7) exemplify progressive verb phrases with *can* and *could*.

(6.6) it's a pity more people *can't be travelling* (ICE-JA, S1A-073)

(6.7) And then there 's the carpark that they *could be hiding* behind different things (ICE-IRL, S1A-038)

An explanation for the low frequency of CAN in general is found in Dixon (2005: 223): he argues that *can* is not used with the imperfective aspect and that *could* is used instead. A closer examination of the present data reveals that the instances of CAN in IrE and SinE are, in fact, *could* instead of the present form *can*, whereas JamE and IndE use *can* and not *could*. Thus JamE especially seems to be using *can* in an innovative way, disregarding the restrictions concerning the use of *can* with the imperfective aspect. Furthermore, CAN actually forms almost 20% of all modal auxiliary use in JamE. Deuber (2014: 221) does not find a similar tendency in the use of *can* and *could* when all verb phrases are considered, on the contrary, their use in JamE is found to largely correspond to BrE. Deuber (ibid.) points out that although *could* is often reported to be used instead of *can* in Caribbean Englishes, this does not apply to JamE to the same extent as it does to Trinidadian English, for instance. Unfortunately, Deuber (2014) does not distinguish the use of modal auxiliaries in progressive or non-progressive verb phrases, but on the basis of the study at hand, it seems that there are differences in the use of modal auxiliaries in the verbal paradigm of JamE.

Taking a closer look at the distribution of modal auxiliaries in two varieties, we see that CAN occurs rather differently in progressive and non-progressive verb phrases. Table 25 below shows the distribution of the modal auxiliaries with progressive and non-progressive verb forms in IrE. The use of CAN is the only major difference between the two (besides the frequency of modal auxiliaries) –

Table 25. Modal auxiliaries in Irish English, progressive vs. non-progressive.

Modal aux.	Progressive	Non-Progressive	Difference in % units
CAN	5.4% (3)	30.1% (453)	-24.7 ***
MAY	8.9% (5)	3.3% (49)	+5.6
MUST	5.4% (3)	3.4% (51)	+2.0
SHALL	5.4% (3)	5.6% (84)	-0.2
WILL	75.0% (42)	57.6% (867)	+17.4 ***
TOTAL	100% (56)	100% (1,504)	

Table 26. Modal auxiliaries in Jamaican English, progressive vs. non-progressive.

Modal aux.	Progressive	Non-Progressive	Difference in % units
CAN	17.1% (6)	40.3% (512)	-23.2 **
MAY	11.4% (4)	4.8% (61)	+6.6
MUST	0	3.2% (40)	-3.2
SHALL	17.1% (6)	8.1% (103)	+9.0
WILL	54.3% (19)	43.5% (552)	+10.8
TOTAL	100% (35)	100% (1,268)	

approximately one third of all non-progressive verb phrases with a modal auxiliary include either *can* or *could* and two thirds either *will* or *would*, while progressive verb phrases with a modal auxiliary most often include *will* or *would*, and on only three instances, *could* was found. This pattern conforms to Dixon's argument on the use of CAN with modal progressives.

Table 26 above, on the other hand, presents the distribution of modal auxiliaries with progressive and non-progressive verb forms in JamE. The situation is quite different from that found in IrE. While the non-progressive verb forms seem to combine with the five modal auxiliaries in a similar manner to IrE (and other varieties, as well), the progressive verb forms in JamE are more evenly distributed than they are in IrE. In JamE, CAN was found in close to 20% of all instances of a modal auxiliary combined with a progressive verb form, which argues against Dixon's statement.

The overall difference in the use of *can* and *could* with progressive and non-progressive verb phrases is clearly seen in Table 27 below. These two modal auxiliaries are ten times as frequent with non-progressives as with progressives.

Table 27. *Can* and *could* co-occurring with the progressive and the non-progressive in World Englishes.

	Modal progressives	Modal non-progressives	Difference in % units
can (have)	2.1% (7)	29.0% (3,377)	-26.9 ***
could (have)	1.5% (5)	8.2% (957)	-6.7 ***
TOTAL	3.7% (12)	37.2% (4,344)	-33.5 ***

CAN is one the most interesting modal auxiliaries to co-occur with the progressives, as the present form *can* should not be used with the progressive (Dixon 2005: 223). The present data, however, reveals that especially JamE acts differently and perhaps innovatively by using *can* with the progressive to a relatively high frequency. Results from Deuber (2014) indicate that further research into modal auxiliaries in progressive and non-progressive verb phrases in JamE may prove fruitful.

MAY - *may* (*have*)/*might* (*have*)

Modal auxiliaries *may* and *might* are found in all but two varieties (AmE and HKE), although their frequency is nevertheless relatively low (in 8.3% of all progressive verb phrases with a modal auxiliary, RF=27). The use of *may* and *might* is illustrated in sentences (6.8) and (6.9) below:

(6.8) I *may be thinking* of shifting because right now I 'm suffering (ICE-PHI, S1A-013)

(6.9) Well she *might be coming* to Clare 's party (ICE-GB, S1A-036)

There is one variety, however, that differs from the others as regards the use of MAY: BrE. In BrE, as much as 25% of all modal auxiliaries co-occurring with the progressive are *may* or *might*, whereas in all other varieties the corresponding figure is less than 11%. However, it must be noted that the overall frequency of modal progressives in BrE is low thus perhaps skewing the picture.

The combination of *may* or *might* and non-progressive verb phrases is also rather rare – in fact, its frequency is even lower than with progressives (see Table 28 below). A possible explanation lies in the data – *may* and *might* are reported to be more frequent in written than spoken data (Leech et al. 2009: 74, 77).

Table 28. *May* and *might* co-occurring with the progressive and the non-progressive in World Englishes.

	Modal progressives	Modal non-progressives	Difference in % units
may (have)	3.7% (12)	2.0% (231)	+1.7
might (have)	4.6% (15)	2.6% (300)	+2.0 *
TOTAL	8.3% (27)	4.6% (531)	+3.7 **

MUST - *must* (*have*)

Must or *must have been* combined with the progressive are used in five of the eight varieties studied, although its overall frequency is low (RF=25). AmE and IndE use MUST most frequently. MUST is used to express epistemic and deontic modality; in

the present data, most instances of MUST co-occurring with the progressive express epistemic rather than deontic modality, as in (6.10):

(6.10) You have you *must be having* waterfalls? (ICE-IND, S1A-008)

When the use of *must* with progressive and non-progressive verbs is compared (Table 29), it can be seen that the situation is almost identical to that with *may* and *might* described above. The frequencies are low throughout, but even lower with non-progressives. Leech et al. (2009: 74, 77) report a decrease in the frequency of non-progressive *must* in both written and spoken registers during the late 20th century.

Table 29. *Must* co-occurring with the progressive and the non-progressive in World Englishes.

	Modal progressives	Modal non-progressives	Difference in % units
must (have)	7.7% (25)	3.0% (348)	+4.7 ***

SHALL - *shall* (*have*)/*should* (*have*)

The overall frequency of SHALL in World Englishes is low (RF=25), but it is found in all varieties studied. To be precise, it is the modal auxiliary *should* which is found in all varieties, whereas *shall* was only found in IndE (two occurrences). The use of *shall* and *should* with the progressive is illustrated below:

(6.11) After ten or fifteen days I *shall be going* there (ICE-IND, S1A-067)

(6.12) It *should be recording* (ICE-HK, S1A-027)

A look at Table 30 below reveals that the use of *shall* and *should* does not seem to relate to the type of the verb pattern it occurs in (progressive vs. non-progressive). The differences found are very small. However, it is clear that *shall* and *should* are used in very different quantities, also with non-progressives. This probably has to do with the type of data used, similarly to *may* and *might* – more instances of *shall* can be expected to be found in data representing more formal spoken situations or written text.

Table 30. *Shall* and *should* co-occurring with the progressive and the non-progressive in World Englishes.

	Modal progressives	Modal non-progressives	Difference in % units
shall (have)	0.6% (2)	0.5% (62)	+0.1
should (have)	7.1% (23)	7.5% (870)	-0.4
TOTAL	7.7% (25)	8.0% (932)	-0.3

WILL - *will* (*have*) / *would* (*have*)

WILL is by far the most frequently found modal auxiliary in the present data (RF=236). The present form *will* is clearly the more frequent form (79.7% *will*, 19.5% *would*, 0.8% *would have*) and the only one found in all eight varieties. Most of the instances of *will* combined with the progressive form found in the data express intention or refer to the future, while *would* is mostly used to refer to future-in-the-past. The use of *will* and *would* with the progressive is illustrated in sentences (6.13) and (6.14).

(6.13) The second six months I *will be tying* up my data (ICE-SIN, S1A-033)

(6.14) See if my settee was making those kind of noises it *'d be going* back (ICE-IRL, S1A-006)

A comparison of the progressive and non-progressive usage of *will* and *would* is seen in Table 31 below. This is a somewhat more complex picture than what the other modal auxiliaries studied yielded. *Will* and *will have* are clearly more frequent with progressives than non-progressives, a situation that on the basis of the data is explained by the fact that *can* is used rarely with progressives, but frequently with the non-progressive. Thus the 'gap' in the paradigm of the progressive with modal auxiliaries left by infrequent use of *can* is filled by *will*. The high frequency of *will* might also be related to expressions of futurity (discussed in section 8.4). On the other hand, *would* and *would have* are, firstly, less frequent than the present tense forms and, secondly, less frequent with the progressive than the non-progressive.

Table 31. *Will* and *would* co-occurring with the progressive and the non-progressive in World Englishes.

	Modal progressives	Modal non-progressives	Difference in % units
will (have)	57.8% (188)	27.3% (3,179)	+30.5 ***
would (have)	14.8% (48)	20.0% (2,331)	-5.2 **
TOTAL	72.6% (236)	47.3% (5,510)	+25.3 ***

To summarise the discussion on the use of the individual modal auxiliaries, the only clear difference in progressive and non-progressive verb phrases regarding modality is that *can* is not used with the progressive, and that *will* is more frequent in progressive than non-progressive phrases. Other modal auxiliaries occur in surprisingly similar quantities as regards the aspect of the sentence they appear in. However, when the use of modal auxiliaries is considered in individual varieties differences are more notable. As noted above, the distribution of the five modal auxiliaries is rather stable in non-progressive verb phrases, while progressive verb phrases show more variation.

Despite the low frequency of the modal progressive, this section has offered an interesting, if brief, insight into modality in combination with the progressive aspect, as well as the non-progressive, in World Englishes. While it has become evident that the use of modal auxiliaries with the non-progressive is stable throughout the varieties, thus indicating an established usage, the situation is not at all as straightforward regarding the modal progressive. Low frequencies throughout the varieties indicate that the combination of modal auxiliaries and the progressive aspect has not been fully accepted. Although the raw frequencies are perhaps too low to make any definite conclusions, it must be remembered that the dataset they were extracted from consisted of 100,000 words and of approximately 13,000 verb phrases per variety. I consider this to be a large enough sample for established patterns of language to be found and recognized in.

To conclude, I shall return to the beginning of the present section, where arguments by Smitterberg (2005) and Smith (2002) on the diachronic development of modal progressives were reported. Based on the present data, it is impossible to make any definite conclusions on whether modal progressives are on the rise or not. A different set of data would be needed to establish a robust picture of what

has happened to the construction in the past decades in World Englishes. However, two possibilities were briefly discussed – that the use of the progressive with modal progressives is indeed decreasing and, on the other hand, that the direction might be the opposite. The declining trend would be led by the traditional varieties, especially BrE and AmE, and would then suggest that the more frequent use of progressives with modal auxiliaries in some of the more recent varieties was an instance of colonial lag. One must consider the level of contact between the varieties at present – does IndE still consider BrE as the source of innovations in language usage, or is it possible that AmE has taken over its place? In either case, the future trend in IndE would be decreasing use of modal progressives. However, it has been argued that IndE might act as a new epicenter for the Asian varieties of English (see e.g. Hundt 2013), thus indicating that IndE might in the future be unaffected by changes in the traditional varieties. Whatever the direction of the change is, it is clear that modal auxiliaries play an important part within progressive verb phrases.

6.3 The perfect progressive

Following Smitterberg (2005: 121), I define the term "perfect progressive" to encompass all progressive verb phrases that include the auxiliary HAVE[112] and a past participle of BE (*have/has/had been Xing*). The combination of the progressive and the perfect results in expressions of "temporary situations or habits leading up to the present with a possible implication that the situation or habit may continue beyond the present" (Quirk et al. 1985: 210ff). Consider (6.15) and (6.16):

(6.15) *Have* you *been doing* much shopping here (ICE-SIN, S1A-003)

(6.16) the rain will probably fall today again because *has been falling* in the afternoon for the past couple of days (ICE-JA, S1A-070)

The use of the perfect progressive in (6.15) entails that shopping has been done on more than one occasion and that this trend is likely to continue in the future. The non-progressive (*have you done much shopping*) would rather entail that the speaker is referring to the current visit to the mall. In (6.16), the unpredictability of rain is

[112] The contracted forms of HAVE are also included.

brought into focus - it is suggested that the rain has not been constant, but rather a set of rainfall followed by drier periods.

Figure 10 below shows the proportion of perfect progressives of all progressives in the data (in the order of frequency). Throughout the present section, it is necessary to take into consideration the low raw frequencies (ranging from 12 to 37), which may affect the reliability of the results.[113]

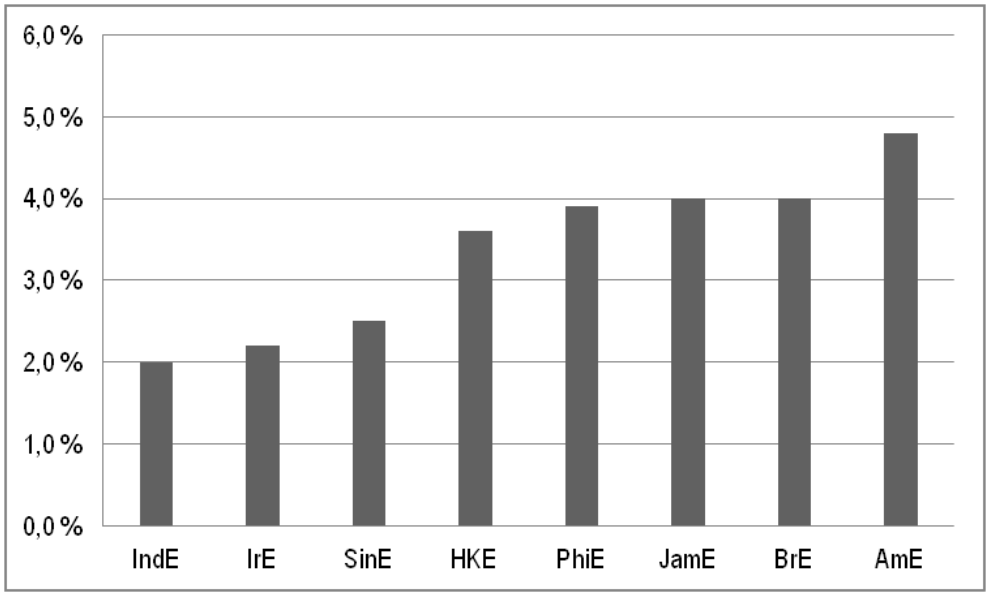

Figure 10. Distribution of perfect progressives in World Englishes.

The lowest frequency of perfect progressives is found in IndE and IrE (2.0% and 2.2%, respectively), and seven of the eight varieties investigated have a proportion of less than 4%. AmE uses perfect progressives clearly more often than any other variety, while BrE and JamE are also high on the scale. While perfect progressives are found in AmE and BrE more often than in most other varieties, the low amount of perfect progressives in IrE stabilizes the situation when comparing IC and OC varieties: there is no statistically significant difference between the two groups ($\chi^2(1) = 0.983$, $p = 0.3215$). Overall, the distribution of perfect progressives across the eight varieties indicates statistically significant differences ($\chi^2(7) = 18.9492$, $p < 0.05$), although the differences between two adjacent varieties (tested

[113] The raw frequencies are presented in Table 32 in the present section.

in the order presented in Figure 10) are statistically non-significant (see Appendix E).

Previous results on the use of the perfect progressive indicate a slight decrease in the frequency of the construction over the recent centuries. According to Smitterberg (2005: 121), in the 19th century perfect progressives accounted for approximately 14% of all progressives, while Smith's late 20th-century data shows a decrease of approximately 50% between the two eras (2002: 319). In her spoken data, Römer (2005: 62) found that less than 5% of all progressives were combined with the perfect aspect, and Collins (2008: 233) reports a frequency of 3.8% for perfect progressives in his spoken data component. This study seems to be well in accordance with the previous studies discussed with 3.4% frequency for the perfect progressive.

A look into the use of perfect forms within the non-progressive paradigm sheds more light on the perfect in general and within the progressive paradigm. As Table 32 below shows, the frequency of perfect non-progressives in the present data is low (3.1%): thus the differences between the frequencies of the perfect progressive and the perfect non-progressive are not great. In fact, the differences are statistically significant in only four of the eight varieties (AmE, JamE, IndE, PhiE). Compared to all verb phrases in the present data, the frequency of the perfect non-progressive in all eight varieties is 2.9% and of the perfect progressive 0.2%. According to Biber et al. (1999: 461), perfect non-progressive verb phrases account for 5-10% of all verb phrases in the four registers[114] included in their study, while the proportion of perfect progressives is less than 0.5% and thus excluded from their analysis. The present results conform to Biber et al.'s findings, albeit being slightly lower.

An interesting feature related to the perfect progressive is the use of temporal adverbials. Kranich (2010: 140) argues that, in BrE, perfect progressives occur regularly with adverbials indicating duration, with the effect that the long duration of the situation is brought into focus. Schlüter (2002) investigates the use of temporal adverbials with the present perfect[115] in five corpora representing BrE and AmE, and finds a very stable picture of approximately a third of all present perfect forms being accompanied by a temporal adverbial. In the present study, approximately 40% of all perfect progressives co-occur with a temporal adverbial,

[114] The registers included in Biber et al. (1999) are conversation, fiction, news and academic.

[115] Schlüter (2002) does not provide information on whether the present perfect progressive is included in the present perfect paradigm of the study.

while the corresponding figure for all progressive forms is less than 20%.[116] Thus perfect progressives are clearly more often modified by temporal adverbials than the progressive overall. Also, compared to the results in Schlüter's study, the perfect progressive is slightly more often temporally specified than the present perfect. These findings together suggest that the perfect progressive is considered a combination that requires overt temporal specification.

Table 32. Perfect aspect with progressive and non-progressive verbs.

	Perfect progressives	Perfect non-progressives	Difference in % units
BrE	4.0% (34)	4.4% (607)	-0.4
IrE	2.2% (24)	2.8% (352)	-0.6
AmE	4.8% (37)	2.3% (318)	+2.5 ***
JamE	4.0% (34)	2.3% (276)	+1.7 **
IndE	2.0% (18)	4.2% (480)	-2.2 **
PhiE	3.9% (31)	2.4% (288)	+1.5 *
SinE	2.5% (18)	2.1% (255)	+0.4
HKE	3.2% (12)	4.1% (412)	-0.9
TOTAL	3.4% (208)	3.1% (2,988)	+0.3

The most frequent type of temporal adverbial co-occurring with the perfect progressive is one indicating duration, *for* [*x amount of time*] (see (6.17) and (6.18) below):

(6.17) Yeah but they *'ve been saying* that <u>for ten years</u> (ICE-GB, S1A-029)

[116] The use of temporal adverbials with the progressive is discussed in section 6.6.3 below.

(6.18) Well they *have been hiding* for most of the day while I was teaching (ICE-HK, S1A-039)

In Schlüter's (2002) data, *for* [*x amount of time*] is the third most common temporal adverbial co-occurring with the present perfect. The most common temporal adverbial in Schlüter's data is *(ever) since (NP)*, which in this study is the second most common temporal adverbial with the perfect progressive.

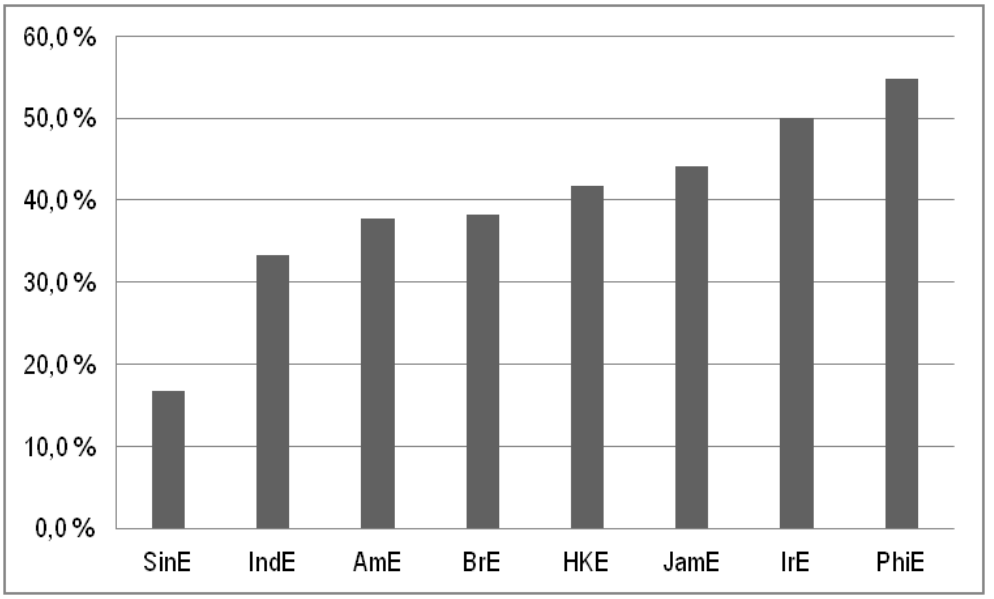

Figure 11. Perfect progressives co-occurring with temporal adverbials.

Majority of the varieties studied show surprisingly consistent amounts of temporal adverbials co-occurring with the perfect progressive (see Figure 11 above). One variety clearly standing out is SinE with a frequency as low as 16%, while at the other end, PhiE has the highest amount of temporal adverbials with perfect progressives (55%). As will be shown in section 6.6.3, PhiE has a high overall percentage of adverbial specification of progressive forms, which may, in part, explain the current situation. However, SinE has a very low percentage of temporal adverbials occurring with perfect progressives, although it shares the top rank with PhiE in the overall adverbial specification. Thus, speakers of SinE may not feel that it is necessary to 'boost' perfect progressives with adverbials of time, while speakers of PhiE feel the opposite. Smitterberg (2005: 188) proposes the following:

> When the progressive becomes more integrated as a marker of ongoing action and imperfective aspect, it can hypothesized that the need for explicit adverbial modification will decrease.

Smitterberg's own data agrees with the hypothesis - the use of temporal adverbials with the progressive decreases diachronically (2005: 190). Unfortunately, Smitterberg does not provide information on how temporal adverbials are used with different forms of the progressive. Nevertheless, based on Smitterberg's hypothesis and the results obtained in this study, it could be argued that the perfect progressive is better established in SinE than it is in PhiE. The use of temporal adverbials and their use with different forms of the progressive is further discussed in section 6.6.3.

Summing up, the perfect progressive is found in each of the World Englishes studied, albeit in low numbers. Highest frequencies were found in AmE and BrE, while the third IC variety, IrE, had a very low frequency for perfect progressives. Results from previous studies form a declining trend for the perfect progressive, and the very low frequencies found in the present study seem to confirm this trend. Very little differences were found in the distribution of the perfect aspect in the progressive and the non-progressive. A tendency to favour perfect progressives rather than perfect non-progressive was found in three of the four varieties with significantly different distributions. Finally, the perfect progressive was found to co-occur with temporal adverbials more often than simple, modal, or passive progressives.

6.4 Voice

The final morphosyntactic variable discussed in the study at hand involves the choice between the active and passive voice. The progressive combines with the passive voice without a change in the meaning of the combination: quite simply the meaning of the progressive (a situation in progress) is merged into that of the passive (less prominence given to the agent) (Biber et al. 1999: 483). The use of the progressive passive is illustrated below:

(6.19) And then she *'s being pursued* by th the New People 's Army (ICE-PHI, S1A-071)

(6.20) Or *are* you still *being messed* around (ICE-IRL, S1A-002)

The progressive passive is known to have been in use since the late 18th century, although it was originally met with prescriptive reaction mainly because the three-verb pattern needed to express it (e.g. *is being built*) was perceived as odd and cumbersome compared to the then prevailing form called the passival (e.g. *is building*) (Hundt 2004a: 79). In present-day English, the progressive passive competes with a number of alternative constructions that are used to express similar situations. These are:

a) the non-progressive passive

b) the *get*-passive, both non-progressive and progressive

c) the active progressive, used transitively with a generalised subject pronoun

Later in this study, I briefly discuss the progressive *get*-passive and the active progressive with a generalised subject pronoun.

The overwhelming majority of progressives (99%) in the present data are in the active voice. Most varieties studied include only a handful of progressive passives (RF=2 to 11), with the exception of JamE and PhiE, where the proportion of passives is slightly higher (RF=18 and 19). Overall, the passive voice is used infrequently in all varieties regardless of their status as an IC or OC variety. This is an expected result since the progressive passive has been shown to be less frequent in the conversation categories compared to the scripted speech or written categories of ICE-GB and BNC (Smith and Rayson 2007: 137).

There does not appear to be a clear patterning of IC and OC varieties as regards the use of the progressive passive (the distribution is statistically significant at $p < 0.05$). As Figure 12 below portrays, the varieties are dispersed rather unevenly: PhiE and JamE have the highest frequency of the progressive passive, with BrE on the third place. In all other varieties, the frequency of the construction is very low (from 0.8% in SinE to 0.5% in IrE, AmE and HKE). However, any attempt at grouping the varieties according to their use of the progressive passive seems somewhat futile as it does not reveal anything important: there is no clear explanation for why IrE and HKE should group together. On the other hand, JamE and PhiE could be argued to group together because of their connection to AmE. This, however, seems to be an implausible explanation for the higher frequency of the passive voice in these two varieties, as AmE itself included only very few instances of the progressive passive. Hundt (2009a: 290) discusses the interesting development of the progressive passive in BrE and AmE over the past decades: while their use in BrE has increased, in AmE it has decreased. The result

of this divergence can also be seen in the present study: the amount of progressive passives is almost three times larger in BrE (1.3%) than in AmE (0.5%).

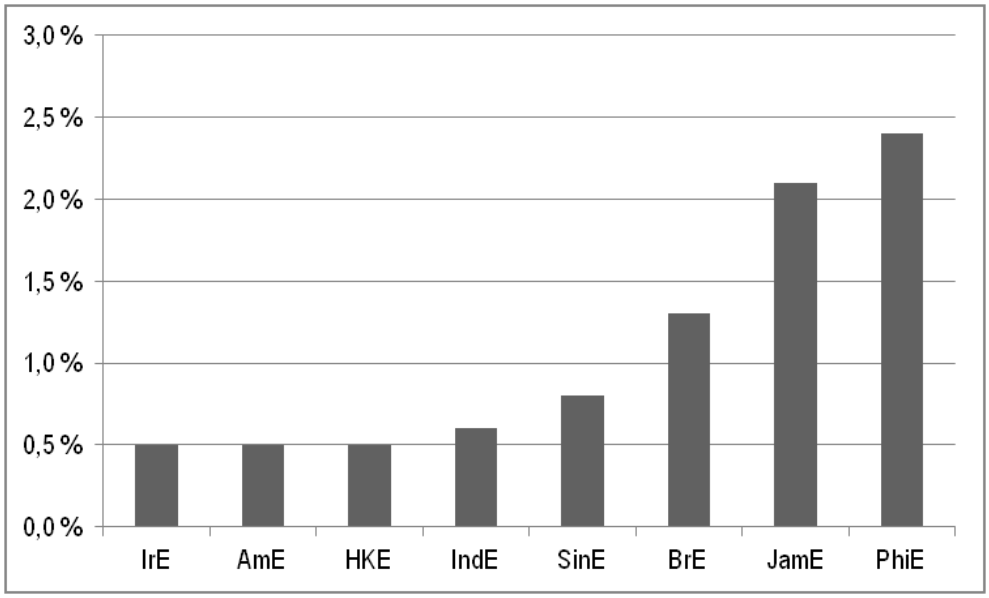

Figure 12. Distribution of the progressive passive in World Englishes.

Smitterberg (2005: 123-33) discusses the historical aspect of the progressive passive. He argues that the development of the progressive *be*-passive is one of the most obvious factors in the spread of the progressive, as it filled a gap in the progressive paradigm. Before the acceptance of BE *being*, the progressive passive was expressed by a construction that formally resembled an active progressive, but was semantically analysed as a passive construction. This construction was labeled 'passival', and is often exemplified in the following way:

(6.21) The house *is building*. passival

(6.22) The house *is being built*. progressive passive

Hundt (2004a) discusses the passival and the emergence of the progressive passive in length. She (p. 105) draws attention to the fact that the increase in the progressive passive coincides with the increase in the overall number of the progressive form, and that the passival had already experienced a decrease in frequency before the progressive passive came into wider use. Smitterberg's (2005: 128) results show that by the end of the 19th century the passival was practically overtaken by BE *being*. However, the passival is not non-existent in present-day

English, although it is very infrequent. Interestingly, possible passival instances were found in the OC varieties in the present data. Thus it could be argued that the passivals in the present data are, in fact, examples of colonial lag. According to Hundt (2004a: 113), the passival in present-day English is strongly related to certain verbs (e.g. PLAY, PRINT, SHIP, SHOW), which is evident in the present data also. The nine instances of the passival appear with four different verbs (SHOW, PRINT, START, WRITE), with SHOW being the clearly most commonly found verb:

(6.23) *Is* Scream still *showing* or that was like last last week already (ICE-PHI, S1A-049)

(6.24) Of course this uh while it (the book) *was printing* the second time I made some addition (ICE-IND, S1A-026)

The passival appears in the present data in such a small proportion that it cannot be regarded as a competitor to the progressive passive. Another competitor relevant to this study mentioned above is the progressive *get*-passive. The progressive *get*-passive, exemplified in (6.25) and (6.26) below, occurs infrequently (RF=17), though evenly distributed in IC and OC varieties (IC 53% : OC 47%).

(6.25) Shetty *is* also *getting promoted* you know (ICE-IND, S1A-058)

(6.26) Walter Ma *is getting marry* on the seventh of this month (ICE-HK, S1A-014)

Similar results are again reported by Hundt (2009a: 301): she found altogether 34 instances which were evenly distributed between IC and OC varieties. Smith and Rayson (2007: 140) point out that the progressive *get*-passive cannot be considered totally equivalent to the *be*-passive, due to the fact that the subject of *get*-passives is usually more actively involved in the realization of the event. This can be seen in example (6.26) above, in which the groom is an essential part of the event of getting married. In (6.25), on the other hand, the subject is not the source of the event taking place.

Smith and Rayson (2007: 141-3) argue that an important alternative to the progressive passive construction is the active progressive used transitively with a generalised subject pronoun (*you*, *we*, *they*). They discuss pronouns *you* and *we* being used with varying degrees of generality: these pronouns may be used to refer to any human being as well as a known individual. The number of active progressives with a generalised *you* or *we* that could be substituted for a progressive passive (see (6.27a) and (6.27b)) has increased in BrE in a period of thirty years (Smith and Rayson 2007: 143).

(6.27a) "The fun of portrait painting," she added, "is in trying to assess and understand the temperament of the people *you are painting*." (LOB C15, cited in Smith and Rayson 2007, 142)

(6.27b) … to assess and understand the temperament of the people who *are being painted*.

The use of generalised *you* and *we* in the active progressive has more than doubled in the period of thirty years studied by Smith and Rayson (2007), which suggests that they are possibly the most important competitor for the progressive passive construction. Within the scope of the present study, however, it does not seem worthwhile to investigate the matter further, as the focus of the study is not on the passive voice itself and as such an investigation would require a considerable amount of manual work. However, the use of generalised *you* and *we* in the active progressive in World Englishes does merit further research.

As already mentioned above, the progressive passive is often considered a key ingredient in the increasing use of the progressive. Mair (2006: 88-95) discusses a closely related matter, i.e. the creation of new progressive constructions, a number of which involve the passive. New complex constructions having been proposed are (ibid.: 90):

a) the present or past perfect passive progressive

I have been being interviewed

b) the future/conditional/modal passive progressive

I will/would/might be being interviewed

c) the future/conditional/modal perfect passive progressive

I will/would/might have been being interviewed

Mair (ibid.) points out that these new constructions are used infrequently, so that there is no direct effect on the frequency of the progressive form on a global level. Indeed, no such complex passive constructions (*be being* or *been being* combined with any number of auxiliaries) were found in the present data. A quick search in some of the complete ICE corpora was conducted in order to find progressive passives in complex environments, but no such instances were found. Thus it seems that the 1-million-word ICE corpora are not large enough to include such a new and scarce construction. Indeed, Hundt (2004a: 99) reports on only one instance of the present perfect passive progressive (*has been being found*) included in the British National Corpus (100 million words).

To summarise the use of the progressive passive in spoken World Englishes, one could argue that although the frequency of the construction is low, there are interesting elements to it. The question remains why the construction is so much more frequent in JamE and PhiE than the other varieties, and why some varieties favour the present tense with the progressive passive so strongly. Possible competitors for the progressive passive were also discussed. The historical form - passival - showed interesting patterning as a possible instance of a colonial lag, while the progressive *get*-passive was evenly distributed in IC and OC varieties thus portraying its stable though infrequent means of expressing the passive voice. Answers to the remaining questions are hard to find in part due to the size of the present data - a feature this infrequent would require a considerably larger set of data in order to bring out the true patterns of usage related to the progressive passive. Also, the type of the present data – spoken rather than written – may have an effect on the frequency of the progressive passive.

6.5 Verb phrase patterns

The four previous sections focused on the use of progressive form as regards tense, co-occurring modal auxiliaries, perfect aspect, and the passive voice, in isolation. In this section, results from the four previous sections are combined to study all the possible verb phrase patterns found in the data. There are twelve possible patterns (exemplified with the verb GIVE in Table 33):

Table 33. Progressive verb phrase patterns.

Progressive form	Active	Passive
Present simple	*is giving*	*is being given*
Past simple	*was giving*	*was being given*
Modal	*might be giving*	*might be being given*
Modal perfect	*might have been giving*	*might have been being given*
Present perfect	*has been giving*	*has been being given*
Past perfect	*had been giving*	*had been being given*

In addition to these twelve verb phrase patterns, fragmentary forms of the progressive were found in the data. These are omitted from the calculations for the time being, and are discussed later in the present section. The distribution of the twelve verb phrase patterns in World Englishes is given in Table 34 below. It is clear that the progressive does not utilize all of its potential in this spoken data on World Englishes: 87% of all progressives occur in two of the twelve possible patterns, active present and past simple progressives. Active progressives, on the whole, account for 99% of all progressives, i.e. the passive voice rarely co-occurs with the progressive in spoken World Englishes.[117] Of the complex combinations, modals (present and past) form the largest group. However, the high frequency of progressives with modal auxiliaries is largely due to the very high frequency of modals in IndE.[118] Present perfect progressives, on the other hand, account for 3% of all progressives in World Englishes. Thus, based on the results discussed above, the spread of the progressive in World Englishes is more likely related to the increasing use of the already established forms rather than to the invention of new

[117] An important factor related to the extremely low number of the progressive passive is undoubtedly related to the spoken data: written data may provide different results. For instance, Collins' (2008: 234) results indicate that the proportion of progressive passives is higher in the written part of his data than in the spoken (5.2% to 1.4%).

[118] The overall frequency of progressives with modal auxiliaries in seven of the varieties studied, excluding IndE, is 3.3%, i.e. close to the frequency of perfect progressives in the data.

combinations. Admittedly, the present data represents spoken language use only; the more complex forms of the progressive are likely to occur more frequently in written language. In the following, the distribution of progressive verb phrase patterns is discussed as regards the individual varieties.

Table 34. Distribution of progressive verb phrase patterns in World Englishes.

Progressive form	Active	Passive
Present simple	60.9% (3,865)	0.9% (57)
Past simple	26.1% (1,659)	0.2% (13)
Modal	4.9% (313)	0
Modal perfect	0.2% (12)	0
Present perfect	3.0% (193)	0
Past perfect	0.2% (15)	0

Table 35 below portrays the distribution of simple, complex and fragmentary progressive verb phrases found in the data. The frequency of simple progressives is overwhelming compared to any other pattern in any variety (c. 87% on average). Also, there is considerable variation between the varieties – the lowest proportion of simple progressives is found in JamE (c. 82%), while the highest proportion is in BrE (c. 91%). IC varieties seem to favour simple progressives, while the OC varieties show more variation as regards simple, complex and fragmentary progressive verb phrases. The varieties investigated may be arranged into three groups according to the morphosyntactic variation attested in them. In the first group, formed by BrE, IrE and AmE, simple progressives are used in roughly 90% of all progressive tokens, while complex progressives (i.e. modal, perfect or passive) account for 8%, and fragmentary forms for the remaining 2%. In the second group, IndE and PhiE, the slightly lower proportion of simple progressives is related to the high frequency of the complex forms (mainly progressives with modal auxiliaries). In the third group, containing JameE, SinE and HKE, complex and fragmentary forms are distributed more evenly, even to the extent that in HKE, the fragmentary forms are more frequent than the complex forms. It thus seems that, in IndE and PhiE, the progressive may indeed have spread through the complex forms: in IndE, modal auxiliaries are clearly the forms that set IndE apart from other varieties. Regarding PhiE, however, it seems that all three complex

forms are more frequent than in most other varieties. In varieties appointed to group three, i.e. JamE, SinE and HKE, it is the fragmentary forms that make the difference – they are used clearly more often than in the other varieties (6% to 9% of all progressives).

Fragmentary forms of the progressive, i.e. progressives in which the auxiliary BE is either missing altogether or incomplete, account for approximately 3% of the present data. Most previous studies exclude the discussion of fragmentary progressive forms, but Römer (2005) does include them in her study on spoken BrE. Approximately 1% of Römer's (2005: 62) data consists of fragmentary progressives – a result which is very close to the present data on spoken BrE (1.3%). As the present data does contain a considerable amount of fragmentary progressives, it seems necessary to discuss them in some detail.

Table 35. Distribution of simple, complex and fragmentary progressive verb phrases.

Variety	Simple	Complex	Fragmentary	TOTAL
BrE	91.0% (766)	7.7% (65)	1.3% (11)	100% (842)
IrE	90.4% (970)	7.9% (85)	1.7% (18)	100% (1,073)
AmE	88.7% (684)	8.3% (64)	3.0% (23)	100% (771)
JamE	81.6% (701)	10.1% (87)	8.3% (71)	100% (859)
IndE	83.7% (745)	14.5% (129)	1.8% (16)	100% (890)
PhiE	87.9% (704)	11.6% (93)	0.6% (5)	100% (802)
SinE	86.8% (636)	7.6% (56)	5.6% (41)	100% (733)
HKE	84.8% (318)	6.4% (24)	8.8% (33)	100% (375)
TOTAL	87.1% (5,524)	9.5% (603)	3.4% (218)	100% (6,345)

Most fragmentary progressives in the present data are simply instances in which the form of BE is missing for some reason. Such cases correspond to approximately 95% of all fragmentary progressives. Sentences (6.28) and (6.29) exemplify a progressive verb phrase where the auxiliary BE is missing altogether:

(6.28) You just *trying* to wind me up (ICE-IRL, S1A-025)

(6.29) You *leaving*? (SBCSAE, 0056

In a few instances, the form of BE is incomplete (*would ø cheering*), uninflected (*we be going*) or otherwise erroneous (consider (6.30)).

(6.30) Where did you *going* to (ICE-HK, S1A-087)

Occasionally, the form of BE is corrupted due to problems in the transcription of the data, as in (6.31):

(6.31) and even so I <unclear>words</unclear> *be going* down there (ICE-JA, S1A-037)

There is considerable variation in the proportion of fragmentary forms between individual varieties: PhiE contains only five (0.6%) fragmentary forms, whereas in HKE the corresponding figure is close to 9%. JamE and SinE also contain a high number of fragmentary forms (c. 8% and 6%, respectively), while other varieties have considerably lower frequencies. The differences between the frequencies of fragmentary progressives in individual varieties may, in part, be explained by differences between the varieties on a greater scale. JamE, for instance, is known for its copula deletion in a number of different contexts (see e.g. Rickford 1999; Deuber 2009), as are HKE (see e.g. Ansaldo 2004) and SinE (see e.g. Wee 2008). Furthermore, Kortmann and Lunkenheimer's (2013) report confirms that the deletion of auxiliary BE before progressives is frequent in SinE and HKE, in particular.

On the basis of the results presented in this chapter so far, we may conclude that the verb phrase patterns in which the progressive is used in World Englishes are largely restricted to the simple progressive, i.e. the occurrence of *am, are, is, was, were* and their contracted forms in a progressive verb phrase. Further, it is the present tense simple progressives that alone account for more than half (61%) of all progressive verb phrases in the present data. Past tense simple progressives are found in approximately every fourth progressive verb phrase. The complex forms

of the progressive are infrequent, which in the present study probably relates to the type of data, i.e. spontaneous dialogues.

Looking at the distribution of progressive verb phrase patterns in IC and OC varieties (see Table 36 below), it becomes evident that the most differences are of a small scale. The slightly higher proportion of simple progressives in the IC varieties is compensated by higher proportions of modal auxiliaries with the progressive, of progressive passives and of fragmentary forms in the OC varieties. The comparison thus emphasizes the fact that IC varieties use perfect progressives more frequently, indicating perhaps that the construction is not fully accepted in OC varieties. On the other hand, progressives with modal auxiliaries are clearly more common in OC varieties, a difference which is explained by the extremely high proportion of such combinations found in IndE.

Table 36. Distribution of simple, complex and fragmentary progressive verb phrases in Inner and Outer Circle Englishes.

	Inner Circle	Outer Circle
SimpProg	90.1% (2,420)	84.8% (3,104)
ModalProg	3.7% (99)	6.2% (226)
PerfProg	3.5% (95)	3.1% (113)
PassProg	0.7% (20)	1.4% (50)
Fragmentary	1.9% (52)	4.5% (166)
TOTAL	100% (2,686)	100% (3,659)

According to the present data, the progressive has spread into World Englishes through the already established, simple combinations. In individual varieties, IndE and PhiE, we find some evidence supporting spread through the increasing use of complex progressive verb phrases.

6.6 Other linguistic parameters

In this section, a number of other linguistic parameters important to the study of the progressives are discussed. The parameters are clause type, contraction and the use of temporal adverbials.

6.6.1 Clause type

The relevance of investigating the clause types that the progressive occurs in arises from the study of the historical aspects of the progressive form. Studies such as Strang (1982) and Smitterberg (2005) indicate that the 19th century is a critical era for the clausal distribution of the progressive. During the 18th century, progressives were usually inserted into subordinate rather than main clauses, while both Strang (1982) and Smitterberg (2005) report an increase in the proportion of main-clause progressives during the 19th century. Kranich (2010) covers a longer period of time, and is thus able to show that, in actual fact, the trend shifted in favour of the main clause already during the 18th century (p. 129). Kranich further points out that main clauses have never been rare during the past four hundred years: the lowest proportion is 38% of all progressive occurrences in the first half of the 18th century. The increasing trend seems to be continuing as Smith's (2002) data reveals an increase in main-clause progressives between the 1960s and 1990s. Turning the attention from historical trends to present-day World Englishes, Collins (2008: 243) reports that main-clause progressives are significantly more frequent than subordinate-clause progressives, and that a slightly higher proportion of main-clause progressives was found in the IC varieties (IC 70% vs. OC 67%). On the basis of the previous studies, main-clause progressives are expected to be more frequent than subordinate-clause progressives in the present study. Of particular interest are the possible differences between the individual varieties, and whether there are traces of the historical developments to be seen.

The progressives in the present data were classified into five different categories: main clause (see (6.32) below), adverbial clause (see (6.33) below), comparative clause (see (6.34) below), nominal clause (see (6.35) below), and relative clause (see (6.36) below), following the classification presented in Carter and McCarthy (2006).

(6.32) Who *'s killing* who (ICE-GB, S1A-049) (main)

(6.33) And she can't call round tomorrow cos I *'m working* (ICE-IRL, S1A-070) (adverbial)

(6.34) it 's better than you *are being* uh raised by the government (ICE-HK, S1A-013) (comparative)

(6.35) Okay I guess it *'s not showing* anymore (ICE-PHI, S1A-049) (nominal)

(6.36) She has two kids and she's the one who *has been trying* to match-make me (ICE-SIN, S1A-023) (relative)

Approximately 1% of the progressives were classified as unclear, because it was difficult to decide how they should be classified. Consider (6.37) below:

(6.37) No I was thinking I *was saving* Ben Elton till after Stephen King but it wasn't (ICE-GB, S1A-048)

In (6.37), the second progressive *was saving* could be classified as subordinate to the first progressive verb phrase, *was thinking*, but it is also possible that the speaker re-formulates his message after the first progressive, and thus the second progressive could be classified as a main clause. Instances such as (6.37) were excluded from the analysis.

The results of the study conform to the expectations (see Table 37 below). In total, 67% of all progressives are found in main clauses, while in the individual varieties, the percentage varies from 59% in JamE to 74% in IndE. IndE and IrE have the highest frequency of main-clause progressives (above 70%), while the lowest frequencies are found in BrE and JamE (below 60%). There does not seem to be any discernible trends regarding the use of main-clause progressives in different varieties around the world. IC Englishes, in particular, show great variation having the lowest and some of the highest frequencies. The distribution of main-clause progressives among the OC Englishes is more even, albeit IndE has a significantly higher frequency than the other three varieties. There is no statistically significant difference between the IC and OC varieties as regards the use of main-clause progressives ($p = 0.0958$).

The subordinate clauses distribute rather unevenly across the four categories employed. Clear majority of subordinate clauses, c. 16%, are nominal (*that*-clauses, *wh*-interrogative clauses, *yes-no* interrogative clauses, and nominal relative clauses), while another large proportion is allocated to adverbial clauses (c. 11% in total). Approximately 5% of all progressives were found in relative clauses, and only a few in comparative clauses. Smitterberg (2005: 196) uses the same classification into five categories, and reports similar results – the order of the clausal categories is the

Table 37. Distribution of the progressive in main and subordinate clauses in World Englishes.

	Main	Adverbial	Comparative	Nominal	Relative	Unclear	TOTAL
BrE	59.9% (504)	16.7% (141)	0.4% (3)	15.2% (128)	7.2% (61)	0.6% (5)	100% (842)
IrE	72.5% (778)	9.2% (99)	0	13.7% (147)	3.2% (34)	1.4% (15)	100% (1,073)
AmE	70.0% (540)	7.7% (59)	0.1% (1)	16.2% (125)	4.3% (33)	1.7% (13)	100% (771)
JamE	58.1% (499)	13.7% (118)	0.1% (1)	20.0% (172)	7.6% (65)	0.5% (4)	100% (859)
IndE	74.4% (662)	7.8% (69)	0	12.7% (113)	3.7% (33)	1.5% (13)	100% (890)
PhiE	66.3% (532)	10.8% (87)	0	18.2% (146)	4.5% (36)	0.1% (1)	100% (802)
SinE	65.8% (482)	12.6% (92)	0.3% (2)	16.8% (123)	4.1% (30)	0.5% (4)	100% (733)
HKE	65.9% (247)	15.7% (59)	0.5% (2)	13.3% (50)	3.2% (12)	1.3% (5)	100% (375)
TOTAL	66.9% (4,244)	11.4% (724)	0.1% (9)	15.8% (1,004)	4.8% (304)	0.9% (60)	100% (6,345)

same, albeit the proportion of main clauses is significantly lower in Smitterberg's 19th-century data. The most significant difference is found in the proportion of relative clauses: in the 19th century, 12% of all progressives were found in relative clauses (5% in the present data). Looking at individual varieties in the present data, the most similar clausal distribution compared to Smitterberg's 19th-century-data is found in JamE, where the proportion of main-clause progressives is 58% and of nominal clauses 20%. It is worth noting that in two varieties (BrE and HKE), the proportion of adverbial clauses is, in fact, slightly higher than that of nominal clauses, which may indicate that another change is on the way – Smitterberg's (2005: 196) results show the proportion of adverbial clauses diminishing across the 19th century, however according to the present results, the adverbial clause seems to have a strong foothold in the clausal distribution of the progressive. Varieties where the proportion of main clauses is high, IndE and IrE, seem to have compensated for the increasing percentage of main clauses by diminishing the use of nominal and relative clauses, in particular.

Considering the present results in the light of previous studies on the clausal distribution of the progressive, a continuing trend towards main-clause usage is established. Smitterberg (2005: 196) reports an increase from 49% to 55% of main clauses during the 19th century, and data from the 21st century give considerably higher percentages: Smith (2002: 325) shows an increase from 56% to 60% between the 1960s and 1990s in BrE, and finally, Collins (2008: 244) indicates that 68% of all progressives in World Englishes were found in main clauses. In fact, one could have expected even a higher main-clause percentage from the present data, as Smith (2002: 326) points out that main clauses being preferred over subordinate clauses is a feature closely related to speech. However, this is true of a number of individual varieties, i.e. IndE, IrE and AmE, where the proportion of main-clause progressives is above 70%.

A further issue to be discussed in relation to the clausal distribution of the progressive is the claim made by Smitterberg (2005: 196) regarding "a positive correlation between (a) the percentage of main-clause progressives in a text and (b) the overall frequency of the progressive". In other words, the higher the main-clause percentage and the higher the frequency of the progressive, the better the progressive may be integrated in a given context, e.g. variety. Smitterberg (2005: 197-202) examined the clausal distribution of the progressive in a number of different genres and was able to establish a pattern according to which genres in which the frequency of the progressive is high, also maintain a high frequency of

main-clause progressives. Table 38 compares the proportion of main-clause progressives to the overall frequency of the progressive in the present data. The varieties are ordered according to the overall frequency of the progressive in a given variety (see column 2), while the third column gives the rank of a given variety according to the proportion of progressives occurring in main clauses (1 = highest proportion).

Table 38. Proportion of main-clause progressives, high to low.

Variety	Rank according to frequency of progressives (RF)	Rank according to frequency of main-clause progressives (%)
IrE	1 (1,073)	2 (73%)
IndE	2 (890)	1 (74%)
JamE	3 (859)	8 (58%)
BrE	4 (842)	7 (60%)
PhiE	5 (802)	4 (66%)
AmE	6 (771)	3 (70%)
SinE	7 (733)	6 (66%)
HKE	8 (375)	5 (66%)

Table 38 indicates that there is a positive correlation between the two features in some varieties studied, while others seem to have a negative correlation. Varieties in which the overall frequency of the progressive is high (IrE, IndE) also manifest a high frequency of main-clause progressives (ranks 2 and 1, respectively), which suggests that the progressive is indeed well integrated into these varieties. Varieties such as PhiE and SinE also seem to correlate positively, indicating a lower degree of integration. However, a negative correlation, i.e. a relatively high overall frequency of the progressive combined with a low frequency of main-clause progressives is found in BrE and JamE. Thus it seems that Smitterberg's claim

cannot be transferred to the present data, although some shared elements are found in the two data sets.

To conclude on the clausal distribution of the progressive, the results of this study conformed to those presented in earlier studies. The proportion of main clauses is significantly higher than that of subordinate clauses, which were further divided into four subcategories. Nominal subordinate clauses were in the majority in all varieties but two, where adverbial subordinate clauses were preferred. Relative clauses and comparative clauses were practically nonexistent in the present data. Looking at individual varieties, it seems unlikely that clausal distribution is critical to the progressive in present-day World Englishes. There are no clear trends as to how the progressive is used in main and subordinate clauses, but rather a freedom of use is manifested. However, varieties where the proportion of main-clause progressives is high also tend to have high overall frequency of the progressive, which indicates a higher degree of integration of the progressive form.

6.6.2 Contraction

In the present study, the following instances of contraction were taken into account when analysing the data: contraction of the verb BE, of the past participle (*have/has/had been*), of modal auxiliaries, or of the negative particle *not* (consider (6.38) to (6.41) below):

(6.38) That 's what they *'re trying* to do I mean (ICE-GB, S1A-003)

(6.39) I*'ve been thinking* about that (SBCSAE, 007)

(6.40) Good morning I*'ll* just *be asking* you general questions (ICE-JA, S1A-085)

(6.41) we *weren't expecting* Mike to be very expressive (ICE-PHI, S1A-059)

Based on previous studies, such as Collins (2008), contracted forms were expected to appear frequently with the progressive in the present data. Collins (2008: 246) reports on more than 40% of all progressives in his spoken data being contracted.

As Table 39 below portrays, approximately 43% of all progressive forms contain a contracted element. Of the individual varieties, BrE and JamE have the largest amounts of contracted forms (55% and 53% contracted, respectively), while SinE (37% contracted) and especially IndE (20% contracted) prefer the full form instead of the contracted form. It is clear that the IC Englishes use more

contracted progressives than the OC Englishes (48% : 40%; $\chi^2(1) = 34.3031$, $p < 0.01$), although PhiE does contain a considerable amount of contraction.

The present results conform well to those presented by Collins (2008: 246). Collins does not separate results from spoken and written data as regards individual varieties, but it is evident that as contraction is a phenomenon mainly found in spoken language (see below); the number of contracted progressives in Collins' study is slightly lower than in this study. The order of varieties studied regarding their rate of contraction is similar to that found in the present study.

Table 39. Contracted progressives in World Englishes.

Variety	Contracted	Full form	Fragmentary	TOTAL
BrE	55.1% (464)	43.6% (367)	1.3% (11)	100% (842)
IrE	42.4% (454)	56.0% (601)	1.7% (18)	100% (1,073)
AmE	47.7% (368)	49.3% (380)	3.0% (23)	100% (771)
JamE	53.3% (458)	38.4% (330)	8.3% (71)	100% (859)
IndE	19.7% (175)	78.5% (699)	1.8% (16)	100% (890)
PhiE	49.1% (394)	50.2% (403)	0.6% (5)	100% (802)
SinE	36.7% (269)	57.7% (423)	5.6% (41)	100% (733)
HKE	39.2% (147)	52.0% (195)	8.8% (33)	100% (375)
TOTAL	43.0% (2,729)	53.5% (3,398)	3.4% (218)	100% (6,345)

According to various previous studies, including Smith (2002), Collins (2008), and Leech et al. (2009), contraction is a sign of spoken language and, more importantly, of colloquialization.[119] The more contracted forms appear in the data the more

[119] Colloquialization refers to "a tendency for written norms to become more informal and move closer to speech" (Leech et al. 2009: 20). See e.g. Levin (2013) for a discussion on the colloquialization of the progressive.

colloquial the variety is. Based on the present data, BrE and JamE are the most colloquial varieties of those included in the study, while HKE, SinE, and IndE in particular are closer to the formal end of the continuum. IC Englishes seem to be more colloquial than those in the OC, while the high rate of contraction in PhiE could be seen as an indication of a colloquializing tendency in the OC varieties. However, a different kind of data would be needed to investigate the matter any further, i.e. a diachronic set of data that would reveal changes that have taken place in recent decades. The very low rate of contraction in IndE (20%) indicates a very low level of colloquialization in IndE. On the other hand, Mair (2009: 52) proposes that in India, English has a status as a formal and prestigious second language, and thus speakers of IndE may consciously avoid using contracted forms.[120] This statement is definitely supported by the present data.

Colloquialization is a phenomenon that is often connected to the question of the spread of the progressive. Smith (2002), for instance, reports on an increase in the number of both contracted forms in general and progressives in general. Furthermore, Smith (ibid.: 327) points out that colloquialization may, in fact, be seen as an explanation for the growing use of the progressive as colloquial language usage, e.g. contraction, spreads into written language. This theory is supported by Leech et al. (2009: 128-9), who report on "massive changes in the use of contractions" which "have an automatic impact on any syntactic pattern involving auxiliaries – including the progressive". The present data shows that contraction is indeed a phenomenon well established in spoken World Englishes, as one would expect when working with spontaneous dialogues, which can be regarded as a good example of colloquial language usage. However, to get an idea of whether contraction is acting as a catalyst of colloquialization and of the spread of the progressive in World Englishes, a different set of data would be necessary (representing a different point in time or a different mode). A glimpse of what is happening is provided by relating the rate of contraction to the frequency of the progressive in a variety – the fact that progressives are more often found in speech than in writing and in more informal rather than formal language situations ties the use of progressives to colloquialization. Thus a high rate of contraction combined with a high frequency of progressives indicates that the variety is more colloquial than one where the rate of contraction and the frequency is lower. Table 40 compares the proportion of contracted progressives to the overall frequency of the

[120] Sharma (2001: 350) discusses the frequent use formal style markers as typical of IndE, and provides as an example a comparison of the frequency of contracted forms in BrE, AmE and IndE.

progressive in the present data. The varieties are ordered according to the overall frequency of the progressive in a given variety (see column 2), while the third column gives the rank of a given variety according to the proportion of contracted progressives (1 = highest proportion).

Table 40. Ranking according to rate of contraction and to frequency.

Variety	Rank according to frequency of progressives (RF)	Rank according to proportion of contracted progressives (%)
IrE	1 (1,073)	5 (42%)
IndE	2 (890)	8 (20%)
JamE	3 (859)	2 (53%)
BrE	4 (842)	1 (55%)
PhiE	5 (802)	3 (49%)
AmE	6 (771)	4 (48%)
SinE	7 (733)	7 (37%)
HKE	8 (375)	6 (39%)

Based on the information at hand, the most colloquial language usage is found in JamE in which both the progressive and contracted forms are frequent. Other varieties towards the more colloquial end of the continuum are PhiE, BrE and IrE, while the least colloquial language usage is found in SinE and HKE. BrE and IrE are interesting in that they are found in the first rank in one of the elements studied, while the other rank is considerably lower – IrE has the highest amount of progressives but uses contracted forms in less than half of all instances (42%), while in BrE, the amount of progressives lies close to the average of all varieties but uses the highest amount of contracted forms (55%). Even more peculiar is the situation in IndE in which the frequency of progressives is high but the rate of contraction is very low compared to any other variety. Thus it is impossible to

judge where IndE should fall on the scale of colloquialization. Although relating the rate of contraction to the frequency of the progressive does provide interesting data, it must be remembered that what they indicate is, at best, tentative. The data is too homogenous for any robust analyses to be performed.

The frequency of contracted forms in progressive verb phrases is relatively high in the present data (43% overall). Only a few varieties stand out: BrE and JamE with their high proportion of contracted forms (more than 50%), and on the other hand, IndE with a very low proportion of contracted forms (less than 20%). JamE, showing a high frequency of contracted forms, presents itself as a colloquial variety, while HKE and SinE may be regarded as leaning towards the more formal end of language use. IndE has been reported to be a prestigious and thus formal language, to which the present results do lend support.

6.6.3 Temporal adverbials

Temporal specification of the progressive has been discussed in numerous previous studies, including Collins (2008) on World Englishes, Römer (2005) on spoken BrE, Smitterberg (2005) on 19th-century BrE, Kranich (2010) on the Modern English period, Smitterberg, Reich & Hahn (2000) on political and academic language, and Scheffer (1975) on BrE. While the frequency of the progressive itself has increased since the 19th century, the amount of temporal specification of the progressive has been reported to have decreased. Smitterberg (2005: 188) argues that increased integration of the progressive form results in a decrease in the need to use temporal adverbials for modification of the progressive form. In other words, the better integrated the form is in a variety, the less temporal adverbials are needed. Kranich (2010: 139), however, rejects Smitterberg's argument, as she shows that the proportion of temporal adverbials (*now* in particular) decreases only slightly during the 20th century. Kranich (ibid.) further points out that the use of the progressive is not always sufficient to specify the temporal location of an event, but may require external temporal specification.[121]

This study follows Crystal's (1966: 8) guidelines according to which any adverbial that gives an answer to the question 'when?' was included in the study, e.g. *at the moment*. Furthermore, following among others Smitterberg (2002: 189),

[121] In present-day English the progressive can express that an event is ongoing at the moment of speaking, e.g. *Outside it is snowing again* (Kranich 2010: 139).

"an adverbial had to both occur in the same sentence as the progressive and be considered to have scope over that progressive in order to be classified as a modifier". The adverbial *for three months* in sentence (6.42) was included in the study as it occurs in the same sentence and has scope over the progressive, while *this summer* in (6.43) was excluded as the temporal adverbial has scope over going to Toulouse rather than the planning:

(6.42) So they *'ve been dating* for three months (ICE-PHI, S1A-081)

(6.43) I *'m planning* to go to Toulouse France this summer (ICE-PHI, S1A-041)

To avoid low frequencies, some multi-word adverbials have been grouped together based on the preposition they include. Thus, following Römer (2005: 77), adverbials such as *for years* and *for half an hour* have been grouped together under the label *for PP* (*for* in a prepositional phrase).

On average 18% of all progressives in the data are modified by temporal adverbials (see Figure 13 below). This proportion is somewhat lower than those reported in earlier studies: 27% in Smitterberg's (2002: 190) historical data, and 24% in spoken BrE (Römer 2005: 75). Collins' study (2008) on World Englishes conforms to the present study: he reports an average of 16% temporal specification in the spoken part of his data (p. 245). In this study, there are significant differences in the frequency of temporal specification of the progressive form in the individual varieties studied. As Figure 13 shows, the scale is from approximately 14% in AmE to as much as 21% in PhiE. Two of the more traditional varieties, BrE and AmE, show the lowest amount of progressives modified by temporal adverbials (c. 15% and c. 14%, respectively), while the highest frequencies are found in PhiE and SinE (c. 21% in both). As regards the distribution of temporal adverbials with the progressive in IC and OC varieties, there is a statistically significant difference ($\chi^2(1) = 11.2497$, $p < 0.001$) indicating that the combination is more often found in OC varieties (IC 16% : OC 19%).

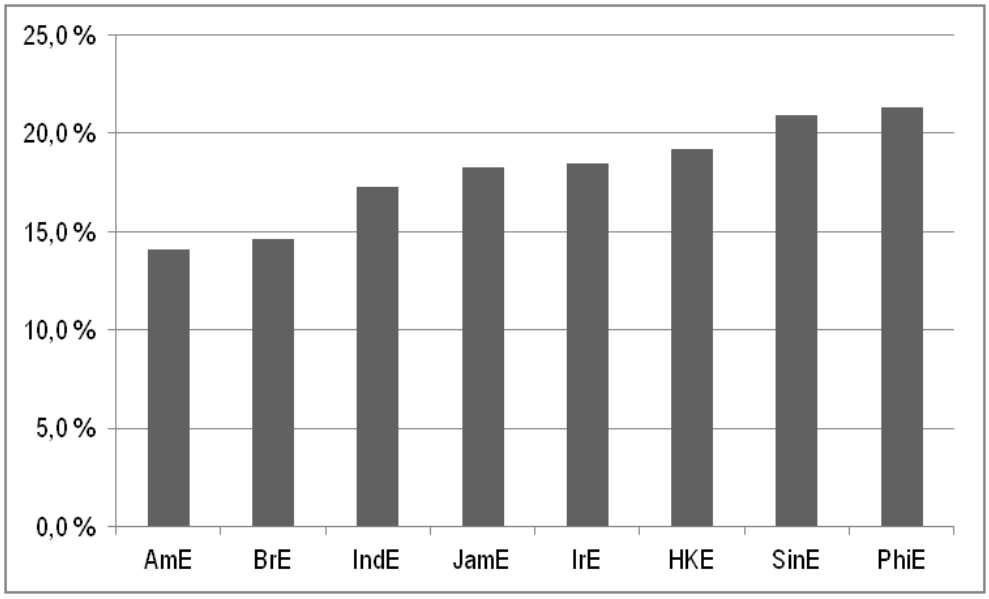

Figure 13. Distribution of temporally specified progressives in World Englishes.

On the basis of the argument made by Smitterberg (2005: 188) according to which increased integration of the progressive form results in a decrease in the need to use temporal adverbials for modification of the progressive form, we may postulate two hypotheses: (1) low percentage of temporal specification of the progressive entails high degree of integration and (2) high percentage of temporal specification of the progressive entails low degree of integration. Looking at the present data from this angle, evidence to support both hypotheses is found. Varieties with low percentage of temporal specification, i.e. BrE and IndE, have mid or high overall frequency of the progressive, i.e. are integrated to a high(er) degree, while varieties with high percentage of temporal specification, i.e. SinE and HKE, have mid or low overall frequency of the progressive. It could thus be argued that the progressive is well integrated into BrE and IndE, while SinE and HKE still require explicit temporal modification. However, the evidence is not straightforward, as the variety with the highest percentage of temporal specification, PhiE, has also a relatively high overall frequency of the progressive. Similarly, on the basis of its low percentage of temporal specification, AmE could be expected to have a high overall frequency, whereas in reality, the overall frequency is relatively low. Thus, while evidence to support Smitterberg's argument does exist in the present data, it must also be noted that the picture is much more complicated. Furthermore, Kranich's view should also be taken into account: the distribution of temporal

adverbials with the progressive may not be related to a *need* to "make the temporal anchoring of the situation more obvious" (2010: 139), but rather to a more general trend in a variety to use temporal adverbials.

The distribution of the five most typical temporal adverbials, or groups of temporal adverbials, is given in Figure 14 below. Some examples of the use of the five temporal adverbials are given in (6.44) to (6.46) below.

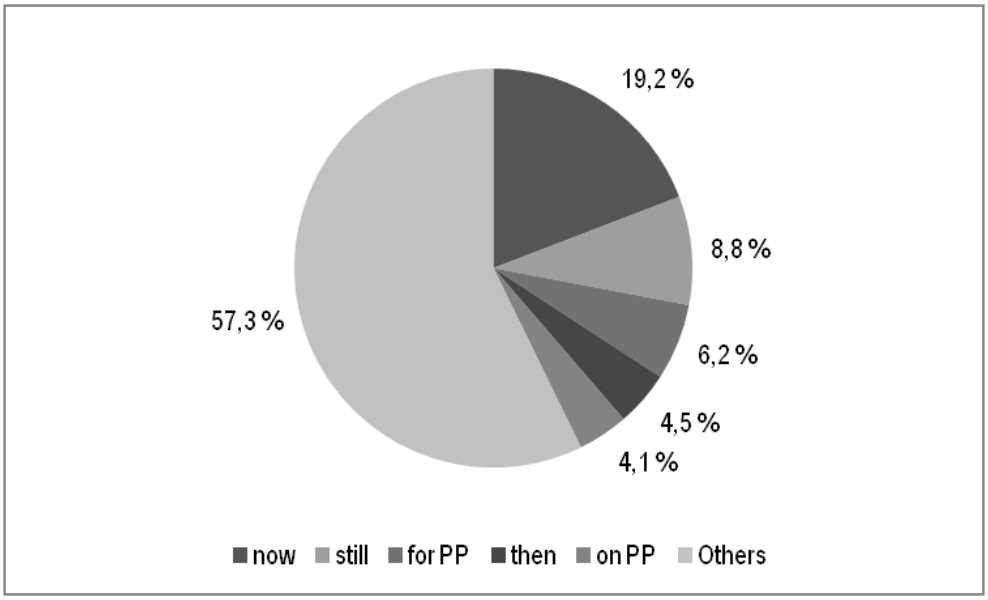

Figure 14. Five most frequent temporal adverbials; percentage of all progressives specified by a temporal adverbial, N=1138.

(6.44) And now I *am doing* my research in one of the topic in English (ICE-IND, S1A-026)

(6.45) *Is* he still *going* out with Estella (ICE-IRL, S1A-016)

(6.46) Well she*'s been* sort of *doing* that for a long long time though. (SBCSAE, sbc0032)

As Figure 14 demonstrates, the most typical temporal adverbials to collocate with the progressive are *now*, *still*, *for PP*, *then*, and *on PP*, which together equal approximately 43% of all temporally modified progressives. By far the most frequent temporal adverbial found in the data was *now*, which is also the single most frequent temporal adverbial in each variety. The popularity of *now* could perhaps be explained by the fact that the progressive is often used to describe

events in progress at the moment of speaking, i.e. *now* (see (6.44) above). Another temporal adverbial that clearly collocates well with the progressive is *still*, which strengthens one of the possible meanings assigned to the progressive – ongoing but unfinished action (see (6.45) above). Considering the five most frequent temporal adverbials in each individual variety, the average proportion for them is 47%. This proportion varies from 35% (IrE) to 68% (HKE) suggesting that the range of possible temporal adverbials found in ICE-IRL, for instance, is greater than that found in ICE-HK. With regard to the proportion of Top5 temporal adverbials in IC and OC varieties, the OC varieties seem to employ a smaller range of temporal adverbials than the IC varieties (Top5 adverbials in IC 41% : OC 51%, $\chi^2(1) = 10.2343$, $p < 0.01$).

Smitterberg, Reich and Hahn (2000), following Scheffer (1975) and Crystal (1966), classify temporal adverbials into three categories: 'frequency', 'duration', and 'time-reference'. Temporal adverbials denoting frequency of occurrence include, for instance, *again* and *every day*, while duration is manifested by temporal adverbials such as *constantly*[122] and *still*. A specific time-reference can be given by temporal adverbials such as *last week*, *tomorrow* or *on Monday*. In the present data, temporal adverbials that give a time-reference are by far the most frequent type in all eight corpora (69% on average, see Figure 15). Approximately 23% of temporal adverbials modifying progressives denote duration, while less than 10% denote frequency. There are only minor differences between the varieties studied regarding the types of temporal adverbials used (no statistical significance). For instance, the proportion of adverbials giving a time-reference is slightly lower in PhiE and AmE[123], and in the same two varieties, the proportion of durational adverbials is higher than in the other varieties.

[122] *Always* and other adverbials with similar meaning are discussed in section 8.2.

[123] Interestingly, PhiE is historically related to AmE: whether the historical connection is the reason for these two varieties showing similar distribution of different types of temporal adverbials is a question of worth asking, however, it is not answered in the present study.

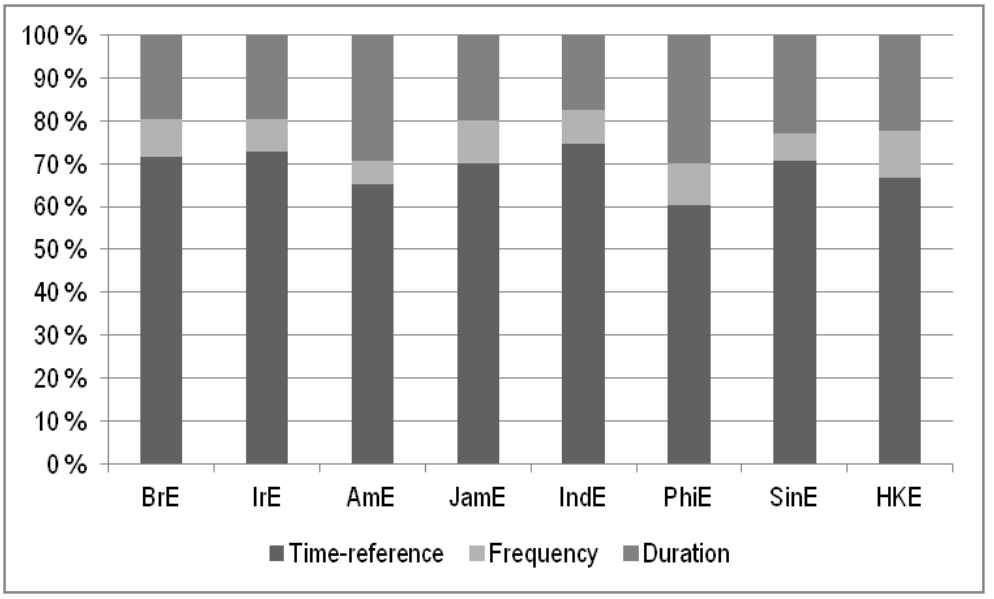

Figure 15. Different types of temporal adverbials modifying progressives in World Englishes.

Examples from each adverbial category are given in (6.47) to (6.49):

(6.47) A lot of people *getting* married this year (time-reference) (ICE-SIN, S1A-039)

(6.48) I *'ve been running* all through the day (duration) (ICE-GB, S1A-065)

(6.49) Yeah they <{> <[> *are* like ... Well </[> *doing* it once a month at most (frequency) (ICE-HK, S1A-089)

Temporally specified progressives were also categorised according to the form of the progressive, i.e. simple, perfect, modal, or passive (fragmentary forms are not included in the discussion; see section 6.5). Examples (6.50) to (6.53) illustrate each of the five categories, and Figure 16 below shows the distribution of temporally specified progressives in each category:

(6.50) Yeah what *were* you *doing* today (Simple) (ICE-GB, S1A-081)

(6.51) They *have been posptoning* his uh promotion right since eleven years (Perfect) (ICE-IND, S1A-058)

(6.52) The second six months I *will be tying* up my data (Modal) (ICE-SIN, S1A-033)

(6.53) So a lot of the personnel *are* still *being* moved around (Passive) (ICE-PHI, S1A-085)

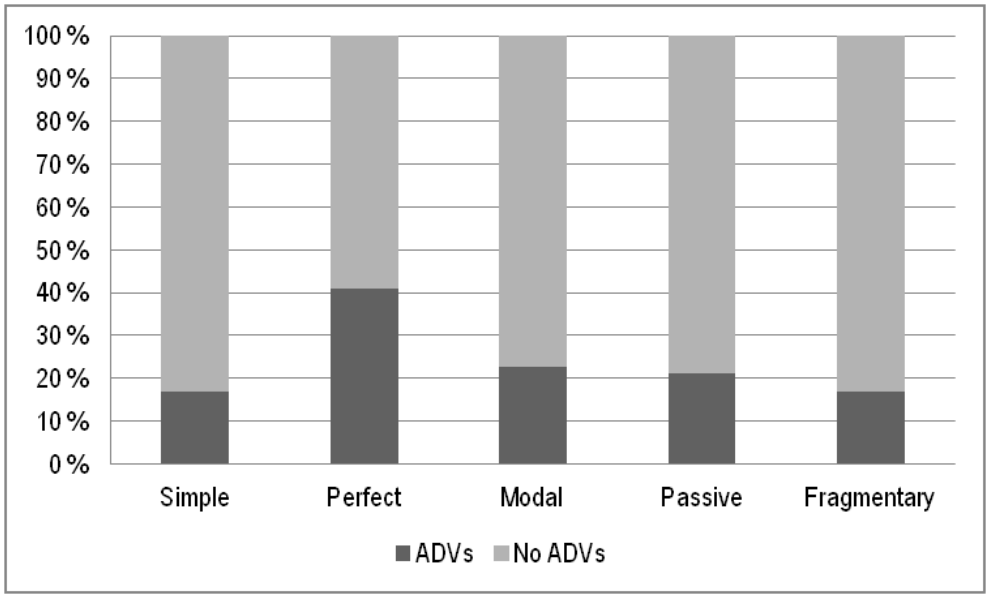

Figure 16. Distribution of different forms of the progressive and temporal adverbials.

As Figure 16 illustrates, simple progressives have the lowest share of temporal adverbials, while 41% of all perfect progressives collocate with temporal adverbials. A possible explanation for the low amount of temporal modification of simple progressives is the fact that the form is well established and thus there is no need for explicit adverbial specification. The opposite may hold for perfect progressives; as the form has not been well established, speakers feel a need to use temporal adverbials. As the raw frequencies for the categories 'perfect', 'modal', and 'passive' are rather low even when considered in total, it is not possible to investigate the varieties individually. However, if the three categories are collapsed into one single category, labelled 'complex progressives', the collapsed raw frequencies may enable comparison of individual varieties. Overall, simple progressives are clearly less often modified by temporal adverbials than complex progressives (c. 17% and 29% respectively). With regard to individual varieties, BrE and AmE have the lowest amount of temporal specification of the simple progressive category, and interestingly, they are among the highest frequencies for temporal specification of complex progressives. This does not support the notion that the progressive form is better integrated into the truly traditional varieties, such as BrE and AmE, than to the more recent varieties. It seems that the raw frequencies are still too low to make any definite conclusions about the integration of the progressive as regards simple and complex progressive categories.

In conclusion, the present section discussed a number of issues related to the temporal specification of the progressive. It was discovered that the more recent varieties, especially PhiE and SinE, use temporal adverbials with the progressive form significantly more often than the more traditional varieties, such as AmE and BrE. These results would suggest that Smitterberg's argument - the better integrated the progressive form is, the fewer temporal adverbials are needed - applies to the temporal specification of the progressive in World Englishes. Based on this hypothesis, results from the present study suggest that the progressive is well integrated into BrE and AmE, perhaps also IndE. On the other hand, PhiE and SinE still require specific temporal modification of the progressive form as it has not integrated into these varieties. Diachronic data would shed light on possible future scenarios, but it is outside the scope of this study as no suitable corpus data exists for the OC varieties, in particular.

Other interesting findings presented in this section were the proportion of the five most typical temporal adverbials to co-occur with the progressive, which suggests that there are differences in the lexical repertoire available to speakers of different varieties, IrE and HKE in particular, and the collocation of temporal adverbials with simple and complex progressives, which suggested that the simple progressive is well integrated into World Englishes, while the perfect progressive in particular seems to still require explicit temporal specification. However, more detailed research into the matter is necessary before any further conclusions can be made. Similarly, as the raw frequencies for the complex patterns were rather low, it was impossible to draw any definite conclusions on the co-occurrence of temporal adverbials and different forms of the progressive in the individual varieties studied.

6.7 Summary of results

Based on results discussed in this chapter, it seems that, in spoken World Englishes, the progressive has spread through the already established forms rather than by extending to new forms to fill in the gaps in the progressive paradigm. The clear majority (87%) of progressives in the present data occur in the category of simple progressives, i.e. *am/are/is/was/were* +V*ing*, while the complex categories, i.e. co-occurrence of the progressive with modal auxiliaries, the perfect, or the passive, are infrequent. As regards tense choices, the present tense is clearly favoured over tha past tense. Individual varieties show moderate variation in the verb phrase

patterns of the progressive. The proportion of simple progressives is clearly higher in the IC varieties, BrE, IrE and AmE, while the complex forms are more frequent in IndE and PhiE than in any other varieties. Finally, fragmentary forms are frequent in JamE, SinE and HKE. IndE and PhiE, in particular, are interesting, as they may be regarded as varieties where the progressive paradigm has indeed undergone some extension: in IndE the progressive occurs very frequently with modal auxiliaries, while all three complex forms are frequent in PhiE. Further research into larger sets of data including both spoken and written material could verify this tentative result of the study.

In addition to the verb phrase patterns of the progressive, the present chapter discussed other parameters of interest: clause type, contraction and temporal adverbials. Generally speaking, the results indicate that progressives in World Englishes are more likely to occur in the main clause (67% in the main clause), in the full rather than the contracted form of BE or *not* (43% contracted), and without a temporal adverbial (18% modified by a temporal adverbial). Individual varieties show a great deal of variation regarding all three features.

Throughout the chapter, reference has been made to the progressive being more or less integrated into an individual variety. This issue arises from Smitterberg (2005), who at length discusses the integration process of the progressive into BrE during the 19th century. Smitterberg (ibid.: 57-8) describes his use of the term 'integration' as related to both grammaticalization and obligatorification, but focusing on areas such as the expansion of the progressive paradigm in order to reach a quantitative point of view. Integration should be regarded as a process in which it is possible that a feature, such as the progressive, is more or less fully integrated into a variety (ibid.), rather than regarding it as an absolute cut-off point. Although the discussion on the process of integration in the present study is not comprehensive, but restricted to only some features included in the study, Table 41 below summarises the results of the integration theory applied to the present data and discussed in various contexts throughout the present chapter. A plus sign indicates a higher level of integration, and a minus sign a lower level of integration. The level of integration is, on most occasions, related to the frequency of a feature in a given variety.

Table 41. Level of integration of the progressive into World Englishes.

	BrE	IrE	AmE	JamE	IndE	PhiE	SinE	HKE
Tense		+	+		-			-
PerfProg	+	-	+		-		-	
Clause type		+			+	-	-	
Temp.ADVs	+		+		+		-	-

Interestingly, there does seem to be pattern emerging, according to which more plus signs are located in IC varieties and more minus signs in OC varieties. Thus according to the information presented in Table 41, the progressive may be better integrated into IC varieties. Of individual varieties, AmE has the highest amount of plus signs, while SinE has the highest amount of minus signs. However, on many occasions contradictory evidence was encountered, i.e. the relevant features yielded mismatched results. It must be borne in mind that, in the present study, the discussion on the integration process of the progressive into World Englishes is best described as a trial rather than robust analysis. Thus the results presented in Table 41 and discussed in the present chapter are tentative at best. Nevertheless, it seems that Smitterberg's discussion could function as a starting point for a well-designed analysis of the process of integration.

On a methodological note, the present chapter compared the morphosyntactic variation of the progressive to that of the non-progressive. These comparisons were a useful tool to bring out interesting facts that would have otherwise gone unnoticed, e.g. the high frequency of modal auxiliaries in progressive verb phrases in PhiE was highlighted by the low frequency in non-progressive verb phrases. On other occasions, the comparison confirmed a tentative idea based on the progressive alone. An example comes from the discussion on tense, where the fact that IrE and IndE behaved differently from the other varieties was confirmed with the help of information from the non-progressive paradigm. Thus it is clear that the progressive vs. non-progressive juxtaposition is a valuable addition to studies on the progressive, especially regarding discussions on the frequency and the morphosyntactic variation of the progressive.

7 Lexical variation

The lexical features included in the present chapter have been shown to be of importance to the development and distribution of the progressive.[124] The focus lies on the main verbs that co-occur with the progressive, their frequency and semantic properties. As pointed out in Chapter 2, some verbs are more compatible with the progressive than are others, due to their inherent semantic properties. For instance, BE and HAVE are rather recent additions to the selection of possible main verbs, while other verbs, such as GO, have for long been established as verbs co-occurring with the progressive. This study provides information on the distribution of frequent progressive main verbs in World Englishes overall, as well as in the eight individual varieties. In addition to investigating progressive main verbs, the present chapter discusses the differences between verbs occurring in the progressive and the non-progressive, as it is possible that a verb prefers to occur in one of them and not the other. The final section of the present chapter discusses the distribution of present and past tense main verbs: Biber et al. (1999: 475) suggest that some progressive main verbs occur more frequently in the past rather than the present tense.

7.1 Frequency of progressive main verbs

In the present data as a whole, a total of 657 different main verbs are used in the progressive. Of these, verbs that occur only once equal 344, verbs that occur two to ten times equal 241, verbs that occur 11 to 25 times equal 31, and verbs that occur 26 to 99 times equal 29. Finally, there are 12 verbs that occur more than 100 times in the data. These 12 verbs alone account for approximately 50% of the data, suggesting that the use of the progressive is heavily concentrated on a small number of verbs. The 20 most frequently occurring progressive main verbs in World Englishes are given in Table 42, in decreasing order of frequency. The most

[124] I would like to thank Professor Douglas Biber for advicing me to investigate the lexical aspects of the progressive and for helping me get started with it.

frequent main verbs occurring in the progressive are GO (9.3% of all progressives), DO (8.5%) and SAY (5.7%), which together account for close to 25% of all instances of the progressive in the present data. As regards the semantic properties of the 20 most frequent progressive main verbs, the majority are activity verbs (e.g. GO, DO) or communication verbs (e.g. SAY, TALK), which are both readily compatible with the progressive.[125] The high frequency of communication verbs may be linked to the increase in the use of the interpretative progressive (see e.g. Kranich 2010; section 8.2 below).

Table 42. The 20 most frequent progressive main verbs in World Englishes.

Rank	Main verb	RF	% (N=6,345)	
1.	*going*	590	9.3%	
2.	*doing*	541	8.5%	
3.	*saying*	359	5.7%	
4.	*talking*	275	4.3%	
5.	*working*	207	3.3%	
6.	*coming*	201	3.2%	
7.	*getting*	195	3.1%	
8.	*trying*	170	2.7%	
9.	*thinking*	167	2.6%	
10.	*looking*	163	2.6%	
11.	*having*	161	2.5%	
12.	*telling*	113	1.8%	
13.	*sitting*	91	1.4%	50% cut-off point
14.	*taking*	90	1.4%	
15.	*studying*	76	1.2%	
16.	*living*	70	1.1%	
17.	*asking*	68	1.1%	
18.	*teaching*	64	1.0%	
19.	*planning*	62	1.0%	
20.	*staying*	55	0.9%	

[125] The semantic domains are further discussed in the following section.

Relating the results from the present study to those presented in Smitterberg (2005: 149) may provide information on how the use of different lexical verbs in the progressive has changed in the recent centuries. What is most notable about Smitterberg's data is the very high frequency of the verb GO (11%) as opposed to any other main verb (second most common main verb is COME with 4%).[126] In the 19th century, the progressive clearly favoured the verb GO, whereas in the late 20th century there seems to be more variety within the use of the most common main verbs. Smitterberg's Table 50 (ibid.: 153) already shows a decrease in the proportion of GO towards the end of the 19th century. The second most common main verb in the present data is DO, which in Smitterberg's data is the third most common main verb to co-occur with the progressive. Looking at the other verbs in Smitterberg's Table 49 (ibid.: 149), we find 13 of the 20 main verbs listed in this study. Close similarities are found also between the current results and those presented by Scheffer (1975: 117). Thus there seems to be a clear continuity of verbs that co-occur frequently with the progressive, starting from the 19th century and continuing to the present day. However, it should be pointed out that Smitterberg's and Scheffer's data represent written language, whereas the present data is spoken. Therefore, some of the differences between the results may be caused by the selection of data, rather than actual diachronical changes in the main verbs co-occurring frequently with the progressive.

The most notable difference between the present results and those obtained by Smitterberg (2005) and Scheffer (1975) is the inclusion of the verb HAVE within the most common verbs to appear in the progressive. Smitterberg (2005: 156-159) discusses the development of BE and HAVE as main verbs in the progressive, as they were added to the progressive paradigm during the 19th century. Smitterberg (ibid.: 156) argues that the addition of BE and HAVE as progressive main verbs affected the general development of the progressive by allowing verbs that usually occurred in stative situations to be used in the progressive in increasing quantities. Smitterberg's data shows a steady increase for HAVE throughout the 19th century, and the emergence of BE towards the end of the 19th century. In the present data, both HAVE and BE are found in all varieties: HAVE is even included within the most common main verbs with a frequency of 2.5% of 6,345 progressives. HAVE is used most often to express a relationship between entities, or to denote an activity (see Collins 2008: 236-7). Consider sentences (7.1) and (7.2):

---

[126] As Smitterberg (2005: 149) points out, the dominance of GO is even more prominent than seems at first, since all instances of *be going to* + infinitive constructions have been excluded from the data.

(7.1) Yeah that time I *was* really *having* a hard time (ICE-PHI, S1A-069)

(7.2) Computer it it plays a great role because we *are having* computers in each field now-a-days (ICE-IND, S1A-061)

The use of *having* in (7.1) is considered standard use of stative verbs in the progressive, whereas in (7.2), *having* refers to a permanent situation rather than a temporary one, and is thus non-standard. See section 8.3 below for further discussion.

BE as a progressive main verb is found in the present data in 0.6% of all progressive verb phrases. Levin (2013: 192-9) discusses the use of the BE *being* + adjective construction at length, and proposes that the construction expresses temporary and specific behavior, rather than a permanent personal trait. For instance, in sentence (7.3) below, the nephew behaves badly on purpose, and it is implied that the rudeness is temporary rather than permanent:

(7.3) and his nephew *was* just *being* rude and obnoxious about every gift (SBCSAE, sbc00044)

Considering the use of these two main verbs in individual varieties, there are a couple of remarks of importance. Regarding the use of HAVE, it is found more often in the OC varieties, or to be precise, in IndE and PhiE (overall distribution in IC and OC varieties yields $\chi^2(1) = 0.3897$, $p = 0.5325$). Within the IC varieties, the progressive main verb HAVE is found clearly more frequently in BrE (RF=27) than in IrE or AmE (RF=18 and 16, respectively), while in the OC varieties there is more variation; HAVE is used very often in IndE (RF=37) compared to any other variety (including the IC varieties), whereas in SinE and HKE its use is infrequent (RF=9 and 5, respectively). Regarding BE, the situation is slightly different. BE, as a progressive main verb, is used in the IC varieties more frequently than in the OC varieties ($\chi^2(1) = 5.2299$, $p < 0.05$), and the two varieties where BE is used as a progressive main verb most often are BrE and AmE (RF=12 and 8, respectively). In the OC varieties, the use of BE as a progressive main verb is very low with raw frequencies of one to three instances per variety. Thus it seems that while HAVE as a progressive main verb has truly gained a foothold in all eight varieties (albeit tentatively in SinE and HKE), BE as a progressive main verb has not accomplished this to a similar extent. Rather, its use seems to be more fully integrated into the two major varieties, BrE and AmE, while the process of acceptance is still underway in the other six varieties. Nevertheless, based on the present data, the emergence of HAVE and BE as progressive main verbs in the 19th century has been

successful, as both verbs are found in all varieties included in the present study. The use of HAVE and BE is further discussed in section 8.3 on stative uses of the progressive.

The five most frequent progressive main verbs in the individual varieties are shown in Table 43 below, along with the proportion of the Top5 main verbs of all main verbs occurring in the progressive. In individual varieties, there is rather little variation regarding the five most frequent main verbs. Verbs GO and DO are found within the five most frequent verbs in all eight varieties, while TALK is found in seven and SAY in six of the varieties. This is not surprising as they are the four most frequent verbs found in the present data overall (cf. Table 42 above). Regarding GO, it is worth noting that its frequency in IrE is significantly higher than in any other variety, and additionally, that it has the highest proportion of all verbs placed in the first rank in the other varieties. IrE on the whole has a higher frequency of the progressive than any other variety (see section 5.2), one of the possible reasons for this being wider range of uses (Filppula 2008). This seems to be true of GO in ICE-IRL: it is used in a considerably more varied way than in the other varieties. A number of instances of GO in IrE may be assigned a communicative reading (see (7.4) below), which is, in the present data, found elsewhere only in BrE (one occurrence). When the special uses are excluded from the data, the frequency of GO in IrE is 10.8%, i.e. similar to rank 1 in other varieties.

(7.4) And he was saying now say right right And she *was going* right right (ICE-IRL, S1A-028)

In JamE, on the other hand, the use of GO is lower than in many other varieties, and it is only the third most frequent main verb in JamE. The two most frequent main verbs, DO and SAY, together account for more than 23% of all main verbs occurring in the progressive in JamE. This proportion is considerably higher than that in any other variety, excluding IrE in which GO and SAY account for a similar proportion. In contrast, the two most frequent main verbs in IndE and PhiE only account for approximately 15% of all progressive main verbs. These facts suggest that JamE and IrE employ fewer main verbs than do IndE and PhiE, or in other words, the distribution of progressive main verbs is more even in IndE and PhiE. Looking at the proportion of the five most frequent main verbs in each variety, the difference between IrE and JamE on the one hand, and IndE and PhiE on the other, is verified. The five main verbs in IrE and JamE account for close to 40% of

Table 43. Top 5 main verbs in individual varieties (% of all main verbs in an individual variety).

	Rank 1	Rank 2	Rank 3	Rank 4	Rank 5	% of Top5
BrE	*doing* (10.3%)	*going* (9.1%)	*saying* (6.2%)	*talking* (4.6%)	*coming* (4.0%)	34.3%
IrE	*going* (15.1%)	*saying* (8.3%)	*doing* (7.7%)	*talking* (4.8%)	*coming* (4.7%)	40.5%
AmE	*doing* (9.9%)	*going* (6.9%)	*talking* (4.8%)	*saying* (3.6%)	*getting* (3.5%)	28.7%
JamE	*doing* (11.8%)	*saying* (11.5%)	*going*(6.3%)	*talking* (3.7%)	*looking* (3.1%)	36.6%
IndE	*doing* (8.3%)	*going* (7.6%)	*working* (5.1%)	*getting* (4.7%)	*having* (4.2%)	29.9%
PhiE	*going* (9.9%)	*doing* (4.7%)	*thinking* (4.2%)	*saying* (4.1%)	*talking* (3.9%)	26.8%
SinE	*going* (10.9%)	*doing* (7.6%)	*saying* (5.2%)	*trying* (4.5%)	*coming* (4.4%)	32.6%
HKE	*studying* (9.3%)	*working* (6.9%)	*doing* (6.9%)	*talking* (6.1%)	*going* (4.5%)	33.9%

all instances of the progressive, whereas in IndE and PhiE, the five verbs only make up for approximately 30% of all instances. Similarly to IndE and PhiE, the proportion of the Top5 is rather low in AmE. To sum up, there is a great deal of variation as to how many different main verbs occur in the progressive in different varieties: some varieties, such as IrE and JamE, are heavily concentrated on a small number of main verbs, while in other varieties, such as IndE, PhiE and AmE, the selection of main verbs is more varied.

The distribution of frequent progressive main verbs is very similar in the first seven varieties listed in Table 43. However, the Top5 in HKE is very different from the other varieties. GO and SAY do not fit into the Top5 in HKE, while the most frequent main verb is STUDY. Another verb not listed in the other varieties is WORK. It seems that the reason for these differences lie in the conversation topics included in ICE-HK. At the time of data retrieval, it was noted that many of the discussions in ICE-HK concern studying and working, thus the verbs denoting these activities must be frequent in the data. However, Schilk & Hammel (2014: 165) point out that speakers of HKE, as well as of IndE, use *studying* to refer to a process of studying and to the state of being a student, especially when a location is mentioned. This stative-like use of the verb *studying* is not found in the IC varieties included in Schilk & Hammel's data (p. 167).

Another difference related to HKE as opposed to the other varieties is the type-token ratio of the data. Although the number of different main verbs is the lowest of all varieties[127], the ratio of different main verbs and the number of progressive verb phrases is the highest in HKE. While the type-token ratio for the progressive in HKE is 0.29, the corresponding ratio in most other varieties is approximately 0.23, and in IrE it is as low as 0.19. Thus lexical variation of the progressive seems to be higher in HKE than in the other varieties, and respectively lower in IrE than in the other varieties. In other words, based on the present data, the speakers of HKE are using a larger set of progressive main verbs to choose from, as opposed to speakers of other varieties, and of IrE in particular. In this light, IrE seems to be lexically constricted as regards the progressive: similarly to JamE, a large proportion of progressive verb phrases in IrE are formed with one (or in the case of JamE, two) main verb. However, it must be noted that the data is too small for any definite conclusions to be drawn.

[127] In the HKE data, there are 110 different progressive main verbs, while the corresponding numbers in the other varieties vary from 174 to 202.

This section discussed the main verbs most frequently co-occurring with the progressive in World Englishes. The verbs GO, DO and SAY were found to be the most frequent progressive main verbs overall, and also in most individual varieties. HKE was found to differ from the other varieties in many respects, most likely due to differences in the topics discussed in the corpora. The following section discusses the semantic domains of progressive main verbs in more detail.

7.2 Semantic domains of progressive main verbs

Biber et al. (1999: 360-5) categorize verbs into different domains according to their semantic characteristics. Some of the domains are more apt to accept the progressive to co-occur with verbs in that particular domain, while other domains reject the progressive. The difference between Biber et al.'s semantic domains and the situation types discussed in Chapter 2 is that the analysis of the semantic domains is based on the lexical meaning of the *verb*, whereas the situation type is determined by the whole *situation* referred to (Smitterberg 2005: 148). The semantic domain of a progressive main verb often functions as a basis for the analysis of the situation type referred to: the two categories are thus at least partially related to each other (ibid.). The situation types of progressive verb phrases were briefly discussed in sections 2.1 above and 8.1 below. In the following, the semantic domains are introduced and followed by a brief discussion of results based on a sample of the present data.

Activity verbs refer to actions and events in which the subject typically has the semantic role of agent (e.g. WORK, COME, TRY). In the present data, the large proportion of activity verbs is partially explained by the fact that the two most frequent verbs, GO and DO, are activity verbs that easily occur with the progressive:

(7.5) Because we*'re going* to his father's house in the morning (ICE-SIN, S1A-007)

(7.6) cos I know they 're all they *'re coming* to Amsterdam and all (ICE-IRL, S1A-038)

Collins (2008: 234) points out that activities that have duration embedded in their core meaning (e.g. SHOP, DANCE) tend to occur more frequently in the progressive than those that do not have duration but rather depict instantaneous actions (e.g. THROW, SHUT).

The other commonly found semantic domain in the present data is *communication verbs*, which are a special subcategory of activity verbs. This domain includes verbs related to speaking and writing (e.g. SAY, TALK, TELL):

(7.6) And I *was saying* it like really loud. (SBCSAE, sbc0036)

Similarly to activity verbs, those verbs that express duration (e.g. CHAT, JOKE) are more common in the progressive than those that do not (e.g. REPLY, THANK).

Verbs that denote a number of different activities and states experienced by humans are called *mental verbs*. Biber et al. (1999: 362-3) list different subcategories of mental verbs: cognitive meanings (e.g. THINK), emotional meanings (e.g. LOVE), perception (e.g. SEE), and receipt of communication (e.g. HEAR).

(7.7) Actually I *am thinking* about tomorrow (ICE-SIN, S1A-049)

Mental verbs may refer to activities that are dynamic (as in (7.8)), or on the other hand, stative (as in (7.9)).

(7.8) What you *studying* right now (ICE-JA, S1A-055)

(7.9) So how *are* you *enjoying* your hostel life? (ICE-IND, S1A-061)

Verbs of existence fall into two subcategories: the first subcategory deals with a state of existence or stance (e.g. EXIST, LIVE, STAY) or with relationships between entities (e.g. CONTAIN, REPRESENT). These verbs are compatible with the progressive, with which they suggest temporariness or limited duration. Consider (7.10) and (7.11):

(7.10) You *were living* together at that time (ICE-JA, S1A-061)

(7.11) well I *'ve been having* problems with my best friend lately (ICE-PHI, S1A-031)

Other verbs of existence concern relationships between entities and descriptions (e.g. BE, SEEM, APPEAR). Such verbs are known as copular verbs, and they are rare with the progressive:

(7.12) We *are being* bitchy (ICE-GB, S1A-023)

Verbs of occurrence report events occurring without volitional activity (e.g. BECOME, DEVELOP). They are compatible with the progressive, as they inherently have duration:

(7.13) No life life *is getting* harder and harder (ICE-HK, S1A-028)

The final two categories are reported to be rare with the progressive (Collins 2008: 236), which is also true of the present data. Verbs of facilitation or causation, or

causative verbs, are verbs that indicate that a person or an inanimate entity brings about a new state of affairs (e.g. CAUSE, HELP, PERMIT):

(7.14) actually they *are not allowing* me to stay in hostel (ICE-IND, S1A-016)

Aspectual verbs describe the stage of progress of an event or activity (e.g. START, CONTINUE, GO ON):

(7.15) I *'m beginning* to hate your ex (ICE-PHI, S1A-057)

The data for the semantic analysis consists of a random sample of 25% of all progressive tokens from each corpus, i.e. 1,590 progressives in total. A smaller set of tokens is easier to manage and thus as accurate an analysis as possible may be reached. Biber et al. (1999: 361) point out that there are two problems related to the classification: first, some verbs cannot be assigned to one single category since their core meanings can actually belong to more than one category at the same time. Such verbs include READ (Communication/Mental) and FIND (Activity/Mental). Second, some verbs may have more than one meanings which belong to more than one semantic domains. Such verbs include CHANGE (Activity/Occurrence) and LOOK (Activity/Mental/Existence). In the present data, such verbs have been assigned to the most likely category based on the context they appear in, or if the context is of no help, the verb has been categorised as 'Indeterminate'. Most verbs, however, have one single core meaning belonging to one of the seven semantic domains.

There are clear semantic preferences that the sample of progressive main verbs attests in the present data. As Figure 16 below portrays, the majority of the progressive main verbs fall into two of the seven semantic categories: more than 50% are activity verbs, while 20% are communication verbs. Approximately 12% of the main verbs were categorized as mental verbs, 8% as existence verbs, and 5% as occurrence verbs. As might have been expected, the proportion of aspectual and causative verbs is practically nonexistent. It thus seems that the restrictions laid out by the semantic domains do apply to progressive main verbs in the present data: domains that are readily compatible with the progressive are in the clear majority, while verbs belonging to semantic domains with clashing semantic properties seldom occur in the progressive.

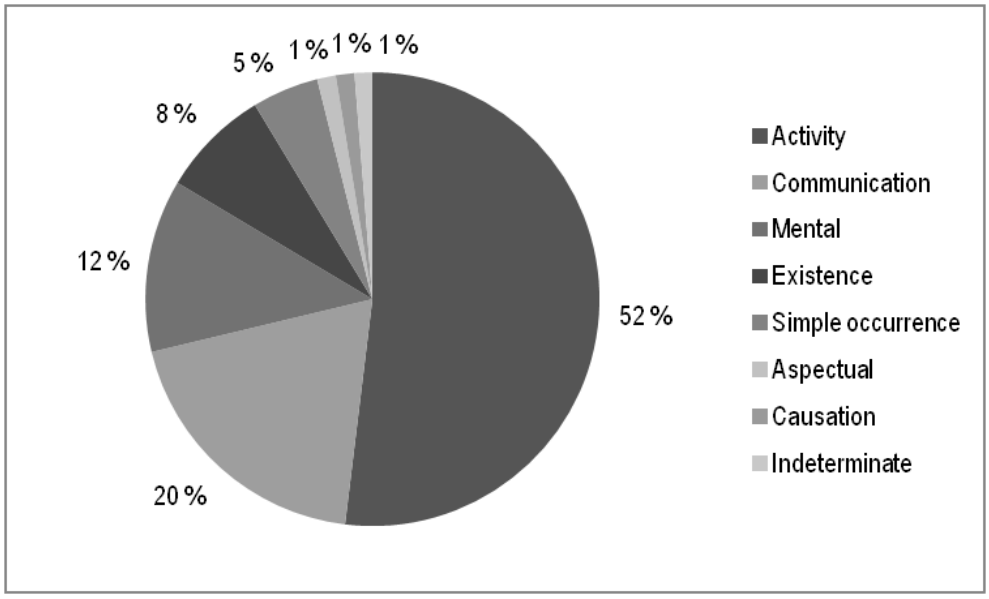

Figure 17. Semantic domains of progressive main verbs in World Englishes.

Smith (2002: 322) shows that in written BrE, the number of activity verbs declined from c. 61% in 1961 to c. 55% in 1991, whereas the amount of communication and mental verbs increased. The present results may be seen as conforming to Smith's findings regarding activity and communication verbs, however, the amount of mental verbs is lower here than might be expected based in Smith's results.[128] Smith (ibid.: 323) links the rise of communication and mental verbs to the increasing use of interpretative progressives. A different point of comparison is offered by Biber et al. (1999: 365) who present the distribution of semantic domains in general (i.e. not only in progressive verb phrases). Compared to these findings, it seems that communication verbs are more frequent with the progressive than they are in general[129], and that causative and aspectual verbs indeed occur rarely in the progressive, as their proportion in Biber et al.'s (ibid.) data is higher than in the present study.

As the number of verbs included in the sample of data is rather low (1,590), it is not worthwhile to look at the distribution of the semantic domains in individual

[128] It must be noted that the present study represents spoken language, whereas Smith (2002) focuses on written language. Thus, the two data sets may not be directly comparable.

[129] According to Biber et al. (1999: 365), the proportion of communication verbs in general is 13%.

varieties. However, the distribution of the semantic domains in IC and OC varieties (see Table 44 below) shows that the differences between the two groups are rather small, but overall statistically significant ($\chi^2(7) = 16.8009$, $p < 0.05$). Activity verbs are more commonly found in the IC varieties, which can be explained by the more frequent use of mental verbs in the OC varieties: of all progressive main verbs in the IC varieties, approximately 9% belong to the domain of mental verbs, while the corresponding figure for the OC varieties is nearly double at 15%. The difference may relate to the stative uses of the progressive, discussed in detail in section 8.3 below.

Table 44. Distribution of the semantic domains in Inner and Outer Circle varieties.

	Inner Circle	Outer Circle
Activity	55.6% (374)	48.9% (448)
Communication	18.9% (127)	19.5% (179)
Mental	8.8% (59)	14.8% (136)
Existence	8.5% (57)	7.2% (66)
Occurrence	4.3% (29)	5.0% (46)
Aspectual	1.2% (8)	1.4% (13)
Causative	1.3% (9)	1.3% (12)
Indeterminate	1.5% (10)	1.9% (17)
TOTAL	100% (673)	100% (917)

To conclude, progressive main verbs in World Englishes most frequently represent activity and communication verbs, which is an expected result, as these are semantic domains that readily co-occur with the progressive. Mental verbs are clearly more frequent in OC Englishes.

7.3 Progressive vs. non-progressive main verbs

Biber et al. (1999: 471-2) discuss lexical associations of the progressive aspect. They list verbs that occur in the progressive more often than in the non-progressive (over 80% or over 50% of the time), and on the other hand, verbs that seldom occur with the progressive aspect (less than 2% of the time). In their data, both dynamic and stative verbs are included in verbs strongly associating with the progressive, but also in verbs rarely occurring in the progressive. Thus the

dynamic/stative dichotomy is not sufficient as a clue for determining whether a verb is frequently used in the progressive or not. Therefore, Biber et al. (ibid.: 473) propose two other features that are useful in distinguishing between verbs that commonly take progressive aspect and verbs that rarely do so. These features are the role of a human subject and the prolongation of the action, state or situation described by the progressive verb. Verbs that frequently occur in the progressive typically have a human subject in the role of an agent, and the action, state or situation described can be prolonged into a process. In contrast, verbs that rarely occur in the progressive take a human subject in the role of an experiencer rather than agent, and they refer to immediate action or to a state that is not a process, i.e. the action, state or situation cannot be prolonged. These two features can be used to distinguish between most verbs either preferring or dispreferring the progressive. For instance, Biber et al. (ibid.) list verbs such as LOOK and LISTEN as frequently occurring in the progressive, while SEE and HEAR are listed as verbs rarely occurring in the progressive. The distinguishing factor between these verbs is the fact that LOOK and LISTEN "occur with a human agent controlling the visual or auditory perception" (cf. (7.18-19) below), whereas SEE and HEAR "describe perceptual states experienced, but not actively controlled" by the human subject (see (7.16-19)[130] below).

(7.16) I was *looking* at that one just now.

(7.17) I felt he wasn't *listening*.

(7.18) I *saw* him the other day.

(7.19) Yeah, I *heard* about that.

The second feature, prolongation of the action, state or situation, is particularly helpful in discussing the numerous dynamic activity verbs that seldom occur in the progressive. According to Biber et al. (ibid.: 474), typical activity verbs, such as DRIVE, are common with progressive aspect because they refer to actions "that can be prolonged over a period of time" (cf. (7.20) below), while many other dynamic

[130] Examples (7.16) to (7.19) come from Biber et al. (1999: 473). Note that SEE can occur in the progressive in the meaning of 'having an ongoing relationship', which according to Biber et al. (ibid.), "is presumably under the control of the subject".

verbs, such as THROW, refer to instantaneous action that cannot be prolonged and thus rarely occur in the progressive (cf. (7.21)[131] below).

(7.20) He was *driving* his van, delivering copies of First Rebel.

(7.21) The man *threw* me off the bus.

Biber et al. (ibid: 471-2) list a large number of verbs that in their data occur either frequently or rarely with the progressive. They give the verbs grouped according to their semantic domain, drawing attention to the fact that most of the verbs with strong lexical preference for the progressive refer to physical or communication verbs, such as SHOP, DANCE, RAIN and SCREAM. However, verbs from the same semantic domain may also prefer the non-progressive, as is the case with ARREST, FIND and REPLY. Mental verbs frequently occurring in the progressive include STUDY and HOPE, while BELIEVE and KNOW are rare with the progressive. Causative verbs are the only domain in which only rarely occurring verbs are listed: as mentioned earlier (in section 7.2), causative verbs are not readily compatible with the progressive. The two characteristics described above, the role of the agent and the prolongation of the action, state or situation, explain the differences between verbs preferring the progressive or the non-progressive.

In the present data, all progressive and non-progressive instances of the 657 main verbs found in progressive verb phrases were searched for with WordSmith Concordancer and the instances were then summed up. This gave a total of 62,500 instances of verbs occurring in either progressive or non-progressive aspect.[132] The proportions of progressive and non-progressive instances of each verb were summed up, and the percentages of progressive and non-progressive instances were calculated. The verbs were also ranked according to the frequency of progressive, non-progressive and all instances. Verbs that were found in the progressive in more than 80% of all instances and were ranked higher than 400[133]

---

[131] Examples (7.20) and (7.21) come from Biber et al. (1999: 474). Note that the verb THROW, among other punctual verbs, can occur in the progressive, but the punctual event is then re-interpreted as a series of repeated acts or the focus shifts to the process leading up to the change of state (see sections 2.3 above and 8.1 below).

[132] This figure does not include the non-progressive instance of the verb BE, as it is somewhat more complicated to capture all of them. In any case, the number of instances of BE is very high making it the most frequent verb in the data. It is very likely that BE prefers the non-progressive as the number of progressive instances of BE is 36 in total.

[133] Most verbs ranking 400 or higher occur five or more times in the data, thus making rank 400 a suitable cut-off point for the present analysis.

(of all instances) were considered to strongly prefer the progressive, while verbs occurring in the progressive in more than 50% of all instances and ranking higher than 400 were considered to prefer the progressive (but not strongly). Verbs that occurred in the progressive in less than 2% of all instances were considered to occur rarely in the progressive. Some verbs (those ranking lower than 400) were excluded from the analysis due to the fact that as many as 137 verbs have a 100% preference for the progressive, but they occur in the data only once (120 of the 137 verbs), twice (13 verbs) or three times (3 verbs). Including these verbs as verbs strongly preferring the progressive would doubtlessly skew the overall results – thus it was decided that the overall frequency of the verb should also matter in the analysis. Table 45 below lists the verbs singled out as strongly preferring the progressive:

Table 45. Main verbs strongly preferring the progressive over the non-progressive in World Englishes.

Main verb	Progs	Non-progs	Total	Rank	proportion of progs
KID	20	0	20	213	100.0
JOKE	14	1	15	241	93.3
WORSHIP	6	1	7	344	85.7
DATE	6	1	7	345	85.7
DEFEND	4	1	5	392	80.0
BRAG	4	1	5	393	80.0

As Table 45 shows, there are six verbs that meet the criteria set for verbs that strongly prefer the progressive - the proportion of progressives of all verb forms is higher than 80% and the overall rank is higher than 400. The verbs do not rank very high on the scale of frequency – two verbs are included within the 250 most frequent main verbs in the data and the other four are clearly less frequent. Indeed, the overall frequencies of the six verbs are not high, ranging from five to twenty. However, the proportion of progressive instances over non-progressive instances

is very high with each main verb. The one verb that has an absolute preference for the progressive in the present data is KID (exemplified in (7.22) below), with 20 instances of the progressive and none of the non-progressive. Another verb with a very strong preference for the progressive is JOKE (as in (7.23) below), with 14 instances of the progressive and one of the non-progressive. What is interesting about these two verbs is that they are both communication verbs and nearly synonymous. They are listed also by Biber et al. (ibid: 471) as communication verbs occurring more than 80% of the time with the progressive.

(7.22) No I 'm just I *'m* just *kidding* (ICE-PHI, S1A-069)

(7.23) and then on the death bed I can say now I *was* only *joking* (ICE-GB, S1A-084)

WORSHIP, DATE, DEFEND and BRAG occur over 80% of the time in the progressive, while none of them are listed by Biber et al. (ibid.). Each of these four verbs has a low frequency overall (five to seven instances), and the progressive instances of WORSHIP and DATE actually come from PhiE. Furthermore, the six progressive occurrences of WORSHIP all come from a single file, while the progressive occurrences of DATE are located in two separate files. Thus the strong preference for the progressive that the present data indicates for these two verbs is based on too narrow a database and so WORSHIP and DATE should be excluded from the discussion henceforth. As regards DEFEND and BRAG, the occurrences come from different corpora and are thus more reliable, albeit low in number. DEFEND is an activity verb (see (7.24) below), and BRAG a communication verb (see (7.25) below).

(7.24) It 's based on a misconception that your party *'s defending* the Union but it blatantly did not (ICE-IRL, S1A-034)

(7.25) They *were bragging* about their sem break (ICE-PHI, S1A-039)

The present data thus confirms Biber et al.'s argument according to which verbs strongly preferring the progressive are verbs of activity or communication (ibid.).

Another group of verbs that also prefers the progressive over the non-progressive, but to a smaller degree, is listed in Table 46 below. The information is presented in descending order according to the proportion of progressive instances, i.e. the verb on the top left is the one with the highest proportion of progressives in this group. The numbers in brackets refer to the number of progressives per all instances (progressive plus non-progressive), i.e. RAIN occurs 22 times in the data, of which 16 are in the progressive.

Table 46. Main verbs preferring the progressive over the non-progressive in World Englishes.

RAIN (16/22)	DONATE (5/8)	UTILISE (3/5)
STARE (6/9)	SUFFER (16/26)	SHAKE (8/14)
MAJOR (6/9)	TEST (8/13)	HEAD (8/14)
LECTURE (4/6)	BACK (3/5)	STRUGGLE (5/9)
RECORD (14/22)	TROUBLE (3/5)	CONTRIBUTE (5/9)
SEARCH (7/11)	SWAY (3/5)	PLAN (62/115)
TEASE (5/8)	CHASE (3/5)	BURN (8/15)

There are 21 verbs included in this group, and similarly to the verbs strongly preferring the progressive, the overall frequency of these verbs is, on most instances, between ranks 250 and 400. An exception is PLAN with 115 instances altogether (of which 62 are progressive), in rank 75. The clear majority of the verbs listed in Table 46 is categorized as activity verbs; among them are included RAIN, SEARCH, CHASE and HEAD, of which RAIN and HEAD are also listed by Biber et al. (ibid.). In addition to activity verbs, LECTURE may be analysed as a communication verb and PLAN as a mental verb. Verbs SUFFER, SHAKE and BURN are more complicated as they seem have more than one possible meaning or semantic domain that they belong to. For instance, the verb BURN may be an activity verb, as in (7.26) below, or an occurrence verb, as in (7.27):

(7.26) they were *burning* these drugs which were out of date (ICE-IRL, S1A-034)

(7.27) The skin is *burning* like whereas in Goa we have the cool climate (ICE-IND, S1A-001)

Interestingly, there is one verb which could be analysed as a causative verb: CONTRIBUTE (as in (7.28) below).

(7.28) I'm resenting this medicine. And I think it's *contributing* to my problems. (SBCSAE, sbc0011)

It must be noted, however, that the frequency for CONTRIBUTE is rather low (nine occurrences), in addition to which the proportional preference is very close to 50%. However, as opposed to the verbs listed as strongly preferring the progressive, the verbs listed in Table 46 represent a larger repertoire of semantic domains, albeit the clear majority of them are activity verbs.

Finally, Table 47 below presents verbs that appear rarely (less than 2% of all occurrences) in the progressive in the present data. Note that verbs that do not appear in the progressive at all are not included in the data.

Table 47. Main verbs preferring the non-progressive over the progressive in World Englishes.

ACCEPT (1/50)	END UP (1/65)	SOUND (1/157)	WANT (4/1805)
LET (4/202)	MIND (1/93)	REMEMBER (3/578)	KNOW (8/3621)
COST (1/53)	SEE (15/1423)	SEEM (1/256)	NEED (1/577)
IMAGINE (2/111)	DECIDE (1/107)	LOVE (1/275)	MEAN (1/788)
CALL (7/420)	UNDERSTAND (2/216)	GUESS (1/309)	
FINISH (2/129)	HEAR (4/433)	LIKE (3/1110)	

The verbs listed in Table 47 are rather different from those listed in Tables 45 and 46. First of all, the overall frequencies of these verbs are considerably higher: most of them are included in the Top100 verbs in the present data. This suggests that verbs preferring the non-progressive are frequent overall, while verbs preferring the progressive, on the other hand, are less frequent overall. Secondly, most of them are mental verbs rather than activity or communication verbs. Verbs that most infrequently occur in the progressive (located in the fourth column of Table 47) include WANT, KNOW, NEED and MEAN (see (7.29) and (7.30) for examples of both progressive and non-progressive usage).

(7.29a) And these women uh village women *are not knowing* anything about it (ICE-IND, S1A-088)

(7.29b) They *don't know* him right (ICE-SIN, S1A-087)

(7.30a) Well we *'ve been meaning* to do it through this friend of mine who 's a photographer for about four years (ICE-GB, S1A-015)

(7.30b) So uhm you *mean* to do the masters over there (ICE-JA, S1A-067)

The progressive in (7.30a) adds duration to the event, as well as an element of 'waxing and waning' (see Huddleston and Pullum 2002: 167), while in (7.29), KNOW is used to refer to a non-delimited stative situation and thus its use differs from that of StE.[134]

The one communication verb included in Table 47 is CALL, and two verbs that refer to activities are MIND and SEE. However, both MIND and SEE have multiple meanings and thus they cannot be categorized as simply activity verbs (see (7.31)[135])

[134] Stative progressives are further discussed in section 8.3.

[135] MIND in (7.31a) is an activity verb, while in (7.31b) it is a mental verb. Interestingly, there are no instances of non-progressive MIND referring to activities.

(7.31a) You *'re minding* the kids (ICE-IRL, S1A-088)

(7.31b) Ya it is true I uh I mean obviously if you want to do a professional presentation then you jolly well *mind* your grammmar (ICE-SIN, S1A-071)

Most of the verbs that rarely occur in the progressive in the present data are mental verbs, such as IMAGINE and UNDERSTAND, while two verbs of existence (SOUND and SEEM) and of causation (LET and COST) are found, and additionally two instances of an aspectual verb (FINISH). Thus there is a clear difference between the semantic domains of verbs that prefer the progressive and those that do not. As mentioned by Biber et al. (1999: 471), verbs referring to physical and communicative activities readily co-occur with the progressive, while mental verbs, in particular, seem to prefer the non-progressive. However, it must be noted that verbs belonging to the same semantic domain may be found on both sides of the line between the progressive and the non-progressive (e.g. communication verbs were found in each of the three groups discussed above).

As regards the individual varieties included in the study, it is difficult to make any definite conclusions on the use of the verbs singled out as preferring the progressive.[136] This is first and foremost due to the fact that as information collected in Tables 45 and 46 are broken down into individual varieties, the raw frequencies of the main verbs become mostly very low, making the analysis tentative in nature. However, a few interesting remarks are discussed in the following. First, five of the verbs listed in Table 46, in addition to the two verbs in Table 45 already discussed, actually only occur in one variety. These are DONATE, TROUBLE and UTILISE in IndE, BACK in BrE and SWAY in PhiE. Similarly to WORSHIP and DATE, these verbs are excluded from the discussion henceforth. The remaining 20 main verbs are more or less equally represented in the individual varieties, although the same verb may prefer the progressive in some varieties and the non-progressive in other. For instance, HEAD is used more frequently in the progressive in IrE, JamE and PhiE, and in the non-progressive in BrE, AmE and IndE. KID and JOKE truly prefer the progressive in all varieties that they are attested in (both verbs are found in five varieties). Other verbs whose place on either Table 45 or 46 is confirmed are RAIN, PLAN, and BURN. Table 48 below

[136] Only main verbs listed in Tables 45 and 46 were analysed as regards individual varieties. Raw frequencies for the progressive instances in Table 47 are too low to make any conclusions. It should also be noted that it is possible that there are, in individual varieties, other verbs fulfilling the criteria set for verbs preferring the progressive: such verbs are not included in the present analysis.

shows also that the nearly synonymous verbs KID and JOKE are dispersed differently in the IC varieties: KID is used in AmE, while BrE and IrE use JOKE. This different pattern of usage is not carried over to the OC varieties. KID was found only in the progressive, while PhiE makes use of both aspects for JOKE. RAIN and PLAN were found in all eight varieties and with both aspects, but the preference is for the progressive in most varieties.

Table 48. Occurrences of progressive and non-progressive main verbs: varieties marked in bold have a preference for the particular aspect.

Main verb	Progressive	Non-progressive	Not attested
KID	**AmE, JamE, PhiE, SinE, HKE**	-	3 varieties
JOKE	**BrE, IrE, JamE, PhiE, HKE**	PhiE	3 varieties
RAIN	**IrE**, AmE, JamE, **IndE**, **PhiE**, **SinE**, HKE	**BrE**, AmE, JamE, IndE, PhiE, HKE	-
PLAN	**BrE**, IrE, **AmE**, JamE, **IndE**, **PhiE**, SinE, HKE	BrE, **IrE**, JamE, IndE, PhiE, **SinE**, **HKE**	-
BURN	**BrE, IrE, IndE, SinE**	BrE, **AmE, JamE, HKE**	1 variety

Although some interesting patterns become visible as the progressive and non-progressive instances of main verbs preferring the progressive are broken down into individual varieties, it must be noted that, on most occasions, the number of occurrences per variety are very low, even one or two. Thus it is impossible to draw any definite conclusions based on the present results. A larger set of data would inarguably bring out the differences or similarities between individual varieties as regards which verbs truly prefer the progressive over the non-progressive.

In addition to the semantic domains of progressive and non-progressive main verbs discussed earlier in the present section, there are other distinguishing factors, which were briefly discussed in the beginning of this section, based on information from Biber et al.'s (1999) study. Whether the results from the present study

conform to Biber et al.'s hypothesis on the role of the agent and the prolongation of the action, state or situation referred to is discussed in the following. According to Smitterberg (2005: 185), the key element in deciding whether or not the subject of a clause acts as the agent is control. If the subject controls the action, state or situation referred to, it can be analysed as the agent of the clause. Consider (7.32) below:

(7.32) they *were burning* these drugs which were out of date (ICE-IRL, S1A-034)

In (7.32), the subject *they* acts as the agent of the clause, as *they* are in control of the burning, i.e. it is possible that they stop burning the drugs. In contrast, the subject of sentence (7.33), *I*, has the role of an experiencer, since the subject does not have control over the situation - it is not possible to stop suffering:

(7.33) I'*m suffering* from cold and fever (ICE-IND, S1A-086)

By determining whether or not the subject has control over the action, state, or situation referred to, it was possible to categorise the subjects of the sentences that had as their main verb one of the 20 verbs listed in Tables 45, 46 and 47.[137] As is evident from Table 49, there is a statistically significant difference between verbs that prefer the progressive and the non-progressive as regards the role of the subject. Within the group of verbs that prefer the progressive, the proportion of subjects in the agent role is more than 70%, while the corresponding figure for verbs that prefer the non-progressive is less than 50%. This result conforms to Biber et al.'s results from a larger set of data from conversation and fiction. Note that non-human subjects were not analysed further, although it is possible for a non-human subject to have control over the situation referred to and thus act as the agent of the clause.

[137] Note that instances of the progressive passive were excluded from the analysis.

Table 49. Role of the subject.

Preferred	Agent	Experiencer	Non-human	Total
Progressive	73.7% (160)	10.1% (22)	16.1% (35)	100% (217)
Non-progressive	45.3% (29)	53.1% (34)	1.5% (1)	100% (64)
Total	67.3% (189)	19.9% (56)	12.8% (36)	100% (281)

The second characteristic proposed by Biber et al. (1999) to influence the choice of progressive or non-progressive aspect is prolongation of the state, situation or action referred to. As was briefly mentioned at the beginning of the present section, the action, situation or state described by some common progressive verbs may be prolonged (e.g. DRIVE), while other verbs, which rarely occur in the progressive, refer to instantaneous action that cannot be prolonged (e.g. THROW). Biber et al. (ibid.: 473) further divide verbs that rarely occur in the progressive into two groups: verbs referring to immediate action, and verbs referring "to a state that is not normally a continuous process". In the present study, progressive main verbs listed in Tables 45, 46, and 47 were analysed into two categories: those that refer to action that can be prolonged, and those that refer to immediate action or to a state that is not normally a continuing process. The hypothesis according to Biber et al. (ibid.) is that verbs belonging to the first category are compatible with the progressive, while verbs belonging to the second category occur rarely in the progressive. Table 50 below shows that Biber et al.'s second characteristic to distinguish between verbs preferring the progressive and the non-progressive also holds true for the present data. Main verbs which prefer the progressive are clearly verbs that refer to actions, situations and states that can be prolonged, while main verbs preferring the non-progressive are more evenly distributed between the two categories.

Table 50. Prolongation of the action, state or situation referred to.

Preferred	Prolonged	Immediate/State	Total
Progressive	87.8% (195)	12.2% (27)	100% (222)
Non-progressive	47.7% (31)	52.3% (34)	100% (65)
Total	78.7% (226)	21.3% (61)	100% (287)

The present section discussed main verbs preferring the progressive or the non-progressive. Main verbs occurring more frequently in the progressive are KID, JOKE, DEFEND and BRAG, while main verb choosing the non-progressive aspect include, among others, IMAGINE, UNDERSTAND, LOVE and KNOW. The semantic domains of the verbs preferring or dispreferring the progressive are mostly distinct: activity and communication verbs clearly prefer the progressive, while verbs preferring the non-progressive are often mental verbs. Although the data set is small, it was possible to find differences between the individual varieties included in this study. For instance, the verbs PLAN and BURN prefer the progressive in some varieties and the non-progressive in others. It was also established that the criteria set up by Biber et al. (1999) regarding the choice of the progressive or the non-progressive hold true for World Englishes, as well.

7.4 Present vs. past tense progressive main verbs

Section 6.1 above discussed the tense choices related to the progressive from a morphosyntactic viewpoint. The present section continues the discussion by investigating whether there are progressive main verbs that tend to prefer the past tense over the present tense. Biber et al. (1999: 475) briefly note that, in their data, main verbs such as *saying* and *thinking* displayed such a tendency, and propose that the past tense use of these verbs convey "a more vivid imagery and a greater sense of involvement". The data here comprises of all progressive instances of main verbs that appear more than ten times in the present data overall, i.e. 74 progressive main verbs with a total of 4,852 occurrences. The tense of each occurrence was noted down per individual variety, after which main verbs that showed a preference for the present tense in each variety were excluded from the investigation. The past tense preference portrayed in the 74 main verbs conforms to the overall morphosyntactic distribution of tense: past tense forms are preferred in AmE and IrE, while IndE and HKE main verbs prefer the present tense. This is an expected result as the amount of main verbs investigated amounts to approximately 75% of all progressives included in the present study.

Of the 74 main verbs investigated, 52 main verbs showed a preference for the past tense in at least one variety, but overall, only seven main verbs occur in more than four varieties and have a proportionate preference for the past tense of more than 50%, i.e. more than 50% of the instances of a given progressive main verb are in the past tense. These main verbs were *crying*, *sitting*, *telling*, *walking*, *watching*, *wearing*

and *wondering*. Neither SAY nor THINK, both mentioned by Biber et al. (1999: 475), are included on the list. Indeed, they both have a preference for the past tense in four varieties, but the percentage for past tense main verbs is slightly less than 50% (47% for both), which is why they are not included in the final set of main verbs preferring the past tense. Table 51 below provides information on the seven main verbs that prefer the past tense in more than 50% of all instances, in more than four varieties.

Table 51. Main verbs preferring the past tense in the progressive in World Englishes.

Main verb	RF	Found in number of varieties	Proportion of past tense progs
crying	22	5	81.8%
sitting	91	7	62.6%
telling	112	6	60.7%
walking	31	4	58.1%
watching	42	6	54.8%
wearing	21	4	61.9%
wondering	38	4	65.8%

As regards the progressive main verb *crying*, its reference is to the past tense in more than 80% of all occurrences. This could relate to the fact it is unnecessary to voice out that one is crying at the moment of speaking, as it usually can be seen or heard by the recipient themselves. Thus most instances of *crying* refer to past events, in which the recipient has or has not been present (consider (7.34)). Interestingly, *crying* is the only one of the seven main verbs investigated here that prefers the past tense with the non-progressive, too.

(7.34) Even when he said sorry he was *crying* then I told him I made him understand see I'm not happy (ICE-IND, S1A-085)

Crying was found to prefer the past tense in five varieties, which are IrE and four of the OC varieties (excluding JamE).[138]

Both present and past instances of *sitting* and *walking* are often used in narratives, as a device to set the situation. Present tense instances of *sitting* are often used for dramatic effect, i.e. are instances of the historic present tense (see Carter and McCarthy 2006: 625) (consider (7.35a) and (7.35b)). The same happens with present tense instances of *walking*.

(7.35a) @you @know I *was* just *sitting* there watching her, telling her ... where to put everything and what not. (SBCSAE, sbc0007)

(7.35b) And I'*m sitting* there with, Lee and Gary's friend, B=- uh, Barney, having a cup of coffee. And all the doors are open, .. into the house @come these two @dogs. (SBCSAE, sbc0032)

Sitting is used in the past tense in as many as seven varieties, although the proportion of past tense forms is not very high, from 55% to 75%. SinE has an equal amount of present and past tense forms of s*itting*. As regards *walking*, three varieties (JamE, PhiE, SinE) have a strong preference for the past tense (83-100%), while AmE and HKE prefer the present tense.

Most past tense instances of progressive *telling* seem to relate to situations where the speaker is recounting what themselves or another person has told at an earlier point in time (see (7.36)).[139] Approximately 20% of the past tense instances of *telling* co-occur with a temporal adverbial with a past reference.

(7.36) I remember this morning Eddie *was telling* me that uhm when the Canadian students leave university they have choice in which jobs they want (ICE-JA, S1A-016)

BrE and JamE strongly prefer the past tense with progressive *telling*, while the past tense preference is more moderate in AmE, IrE, PhiE and SinE. In IndE and HKE, the present tense is preferred with *telling*.

As regards *watching* and *wearing*, there does not seem to be a great difference in the use of present and past tense instances. For *watching*, the proportion of past

[138] The raw frequencies of the progressive main verbs under investigation are rather low (except for *sitting* and *telling*), making the results on individual varieties tentative in nature.

[139] Such use of the progressive may be labelled 'recentness progressive', and is characterized by the co-occurrence of a past tense progressive with a temporal adverbial such as *just* (Pfaff, Bergs and Hoffmann 2013: 221).

tense instances is approximately 55% indicating that the preference for the past tense use in the progressive in the present data is not great. Past tense *watching* is strongly preferred in AmE, JamE and HKE, whereas IndE has an equal amount of present and past tense forms, and SinE prefers the present tense. As regards *wearing*, the situation is even more tentative, as the past tense is rather moderately preferred in four varieties (60-80%), and the present tense is used more often in two varieties (JamE, IndE).

The difference between present and past tense instances of the verb *wondering* is not due to difference in meaning, but rather to difference in pragmatics. The past tense form of *wondering* is often felt to be more polite than the present tense form, and is often accompanied by words such as *just* or *only* (see e.g. Römer 2005: 126, 164-5):

(7.37) I *was* just *wondering* if it was worth complaining to whoever was in charge or not bothering (ICE-GB, S1A-069)

The past tense forms of *wondering* are in the majority in four varieties (BrE, AmE, IrE and SinE), whereas PhiE favours the present tense forms of *wondering*.[140]

As regards the use of the seven progressive main verbs in individual varieties, it turns out that not all varieties show preference for the past tense. This is the case with IndE, for instance, in which less than 40% of all the instances of the seven verbs are in the past tense. There are a few other verbs that do prefer the past tense within the 74 main verbs occurring more than 10 times in the present data, but these verbs did not meet the overall criteria set for the present analysis. Most other varieties do prefer the past tense with the seven verbs singled out, the proportion of past tense instances of these verbs being more than 60%, or more than 70% in the case of BrE.

Overall, there is rather a great deal of variation in the distribution of progressive main verbs preferring the past tense, regarding both the varieties in which the verbs occur and the proportion of past tense instances. However, the seven verbs do meet the criteria set, i.e. they occur in more than four varieties and the proportion of past tense instances in more than 50%. Moreover, on most occasions the use of the past tense is related to the function of the progressive, e.g. tentativeness in the case of *wondering*, or narratives with *sitting* and *walking*. These uses of the progressive differ somewhat from the prototypical progressive meaning

[140] *Wondering* is further discussed in section 8.2 below.

in that they do not necessarily imply duration or ongoingness. Indeed, they are more suitably analysed as subjective progressives, which are further discussed in Chapter 8.

7.5 Summary of results

The present chapter has shown that the level of lexical variation regarding the progressive in World Englishes is ultimately rather low, i.e. there are not many significant differences in the use of progressive main verbs in the World Englishes investigated. The most frequent progressive main verbs – GO, DO and SAY – are shared in seven of the eight varieties studied; HKE has a very different distribution of frequent main verbs, but this is most likely due to the topics of the conversations included in the present data.[141] In diachrony, by comparison of the present and previous studies, a continuity of frequent progressive main verbs was established from the 19th century to the present day. One of the most important aspects related to progressive main verbs is the occurrence of the verbs BE and HAVE: the present study showed that these two verbs are used in Englishes throughout the world, albeit in different numbers. The distribution of HAVE is more stable, indicating its establishment as a common progressive main verb. Regarding BE, it seems that the verb is still undergoing the process of integration, as it is found mainly in BrE and AmE, and only infrequently in the OC varieties. The importance of integrating BE and HAVE into the progressive paradigm is related to the breakdown (or rather, loosening) of the semantic restrictions traditionally regarding the use of the progressive. Allowing these, and other, stative verbs to be used in the progressive contributed to the increasing use of the progressive overall (Smitterberg 2005).

The investigation into main verbs that prefer to occur in the progressive rather than the non-progressive revealed a fundamental difference in the semantic domains that these verbs represent. Main verbs occurring frequently in the progressive are mostly activity or communication verbs, whereas verbs occurring rarely in the progressive are, on most occasions, mental verbs. The high frequency of communication verbs may be related to the increasing use of interpretative progressives and to the frequent use of narrative progressives in speech (see e.g.

[141] In order to either confirm or reject the finding, an investigation into a larger set of data is needed. Such an investigation is beyond the scope of the present study.

Kranich 2010). Another connecting factor between subjective uses of the progressive and the lexical variation attested in progressive main verbs lies in the use of past progressives: past tense progressive main verbs seem to be related to expressing tentativeness (*wondering*), or to setting the scene in narratives (*sitting, walking*).

8 Functions of the progressive in World Englishes

One of the most interesting aspects related to the progressive in World Englishes are the functions in which the progressive is used in different localities around the world. A number of previous studies and surveys (e.g. Hundt and Vogel 2011, van Rooy 2014; Kortmann and Lunkenheimer 2013) indicate that variation is found especially within stative uses of the progressive in varieties such as IndE and BSAfE. What most previous studies concentrate on are the so-called extended uses of the progressive, i.e. the use of the progressive to refer to non-delimited stative or habitual situations. The present study, however, provides information on the full spectrum of usage, including the proportion of prototypical progressives, stative and habitual uses of the progressive, futurate progressives and subjective progressives, in all eight World Englishes included.

Section 5.1 above made reference to a number of possible explanations for the increasing frequency of the progressive – some of them are related to the functions that the progressive is used in. Both Mair (2006) and Smith (2005) propose that the progressive is increasingly being used with stative verbs; sections 7.2 and 7.3 above already briefly discussed the use of stative verbs in the progressive, and the present chapter takes a more detailed look at stative progressives. While Mair (2006) discusses the increasing frequency of *already established* uses of the progressive, Smith (2005) focuses on the spread of the progressive to *new* uses, e.g. uses that are more subjective than the prototypical functions. Although this study is synchronic rather than diachronic, and thus unable to provide information on the diachronic development of functions of the progressive in World Englishes, it is possible to contrast IC and OC varieties and to investigate whether there are major differences between the two groups. Of particular interest are stative, habitual and subjective uses of the progressive. Many grammar books state that the progressive should not be used with stative and habitual situations, as the semantic properties of the progressive clash with those of statives and habituals. On the other hand, IndE for instance is often characterized by extension of the progressive to stative and habitual situations (see e.g. Balasubramanian 2009). Subjective progressives, in turn, are an example of the progressive acquiring new functions: Kranich (2010) reports on developments related to interpretative progressives, a subtype of subjective

progressives, in Modern English and suggests that they could, in fact, be considered a catalyst for the increasing frequency of the progressive in general.

The data used in the present chapter consists of the first 250 progressives in each component of the ICE included in the study, i.e. 2,000 progressives in total. Fragmentary and perfect progressives were excluded from the data in order to avoid instances where the analysis may run into problems (see section 4.1.4 above). Each progressive was analysed separately taking into account as much of the context surrounding it as was necessary to be able to determine the function.[142]

Table 52. Functions of the progressive in present-day English.

Function	Subgroups	Code
Progressive		PROG
Stative	1) Delimited stative (standard)	PSTA
	2) Non-delimited stative (extended)	ESTA
Habitual	1) Delimited habitual (standard)	PHAB
	2) Non-delimited habitual (extended)	EHAB
Subjective	e.g. attitude, emotion, tentative, interpretative	SUB
Futurate	1) *will*/*shall* + *be* V*ing*	FUT1
	2) Futurity evident from context	FUT2
Other	(e.g. WORK-verbs)	OTH
Indeterminate		IND

[142] I am indebted to Dr. Heli Paulasto and Dr. Lea Meriläinen at the University of Eastern Finland for their help with analyzing a part of the present data. Approximately 50% of the present data is used for the analysis in Meriläinen et al. (fc.).

Each progressive was assigned one function, although, in some cases, two functions could be argued to overlap. Table 2 is reproduced opposite as Table 52, showing the categorization of the functions. For an introduction of the functions of the progressive in present-day English, see section 2.3 above.

Figure 18 below gives the overall distribution of the functions of the progressive in World Englishes. Approximately 39% of the 2,000 progressives investigated refer to prototypically progressive situations, i.e. ongoing and/or incomplete situations. The largest group of progressives which carry a meaning other than the prototypical progressive one is, quite interestingly, the subjective progressive (17.5%). This finding lends support to Kranich's argument on the importance of subjective progressives as regards the increasing frequency of the progressive in general. Stative progressives form the third group with 14.5%, and approximately 12% of all progressives refer to situations in the future. Habitual progressives are found in 8% of all instances – the low figure conforms to Kortmann and Lunkenheimer's (2013) report. The category labelled 'Other uses' consists mainly of WORK-verbs, i.e. verbs such as *working, teaching* or *studying,* which

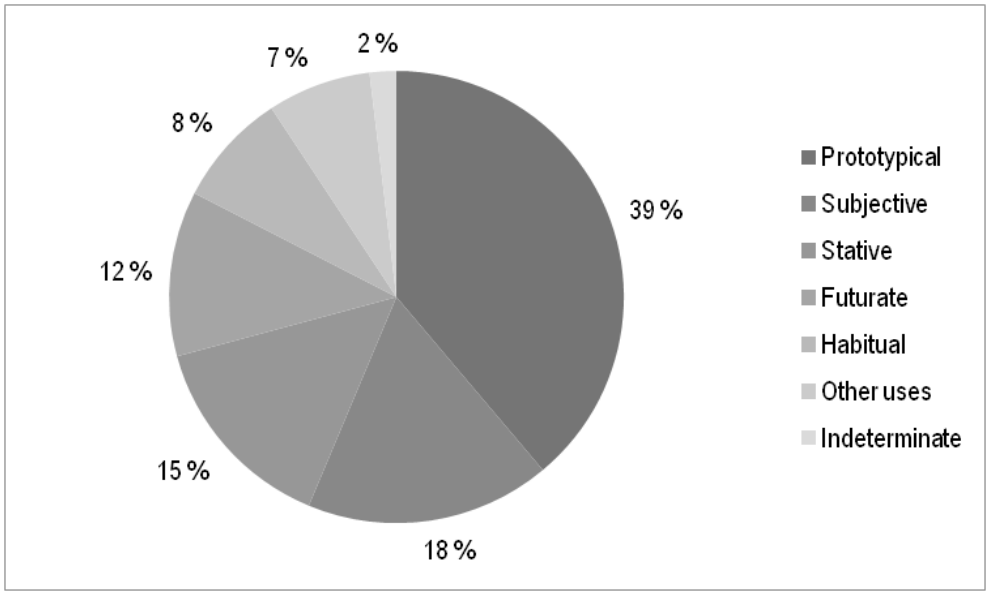

Figure 18. Functions of the progressive in World Englishes.

are difficult to assign to one category only as they may refer to progressive or habitual situations (consider (8.1) below).[143] A small amount of progressives in the present data were impossible to categorise, mainly due to incomplete transcription (as in (8.2) below) or due to difficulties in choosing only one category (as in (8.3) below which could be analysed as a futurate, habitual and possibly also a prototypical progressive).

(8.1) What they 're doing is they *'re working* on the Pascal thing (ICE-GB, S1A-005)

(8.2) And we'*re moving* <unclear>words</unclear> university (ICE-JA, S1A-020)

(8.3) So where *are* you *putting* up now (ICE-SIN, S1A-035)

Some interesting differences arise when the functions of the progressive are grouped together according to IC and OC varieties (see Figure 19 below). The prototypically progressive uses are slightly more frequent in IC varieties, which indicates that the more 'special' uses are more frequent in OC varieties. However, it seems that not all 'special' uses act similarly: it is rather the use of futurate progressives and of WORK-verbs that are more frequent in the OC varieties. In contrast, subjective and stative progressives are more frequent in the IC varieties, while habitual progressives are distributed very evenly. The overall distribution of the functions of the progressive in IC and OC varieties is statistically highly significant ($\chi^2(5) = 33.1214$, $p < 0.001$). However, investigating the functions individually, we find that only within the use of subjective progressives do IC and OC varieties distribute in a way that is statistically significant ($\chi^2(1) = 4.1457$, $p < 0.05$).

Another statistically significant difference is found in the distribution of the so-called WORK-verbs ($\chi^2(1) = 26.6373$, $p < 0.001$). The large proportion of WORK-verbs in the OC varieties (HKE and IndE in particular) may, in fact, distort the results regarding the distribution of the 'proper' functions of the progressive. If 50% of the WORK-verbs in each corpus are re-analysed as prototypical progressives and the other 50% as habitual progressives, it becomes evident that all of the 'special' uses of the progressive differ as regards IC and OC varieties. The re-analysis of the WORK-verbs diminishes the gap between IC and OC varieties within

[143] WORK-verbs are most frequent in HKE (as much as 19% of all progressives analysed) and IndE (14%); for the other varieties the frequency of WORK-verbs is approximately 4%. As discussed in section 7.1, *working* and *studying* are the most frequent progressive main verbs in HKE, most probably due to the topics of conversations included in ICE-HK.

the category of prototypical progressives, while within the habitual category, the OC varieties would then be more frequent than the IC varieties (IC 10% : OC 13%).

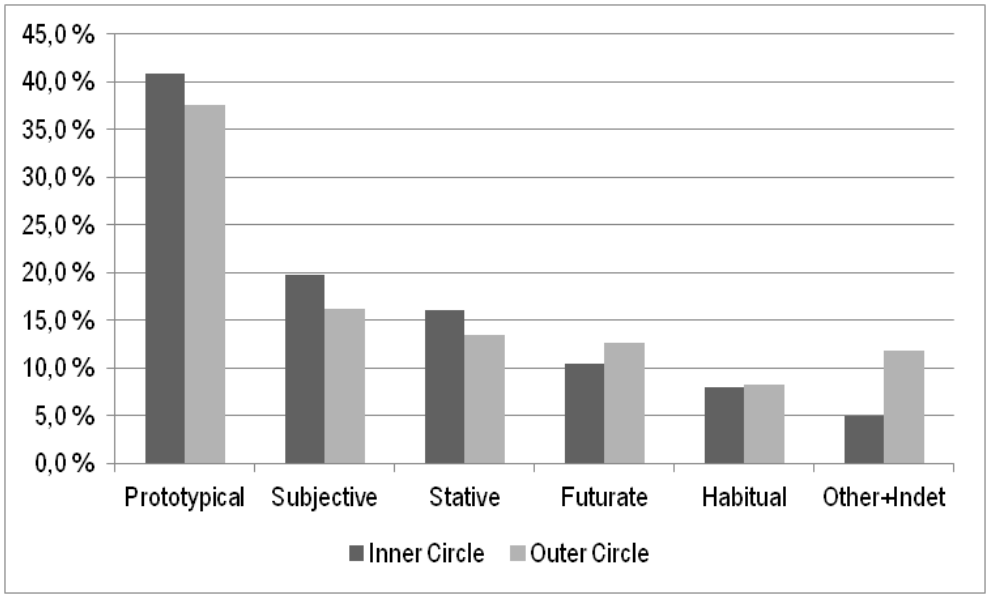

Figure 19. Functions of the progressive in Inner and Outer Circle Englishes.

Interestingly, stative progressives have often been attributed to OC Englishes (see e.g. Kortmann and Lunkenheimer 2013), but here it is the IC varieties with more frequent use of stative progressives. However, a different picture may arise when the categories are investigated in more detail. Indeed, one of the most interesting factors regarding the functions of the progressive in World Englishes is the use of the non-standard or extended uses of stative and habitual progressives (discussed in sections 8.3 and 8.5 below), as well as the use of the different subjective progressives (discussed in section 8.2 below). Futurate uses of the progressive are discussed in section 8.4 below.

8.1 Prototypical progressives

The largest group of progressives (39%) in the present data refers to situations that are portrayed as ongoing and incomplete at the time of utterance, i.e. are used in the prototypical function of the progressive.[144] As Figure 19 above shows, prototypical uses of the progressive are slightly more frequent in the IC varieties (IC 41% : OC 38%, p = 0.1387)). The fact that prototypical progressives are more common in IC varieties suggests that OC varieties may indeed be characterizes as having a wider range of uses.

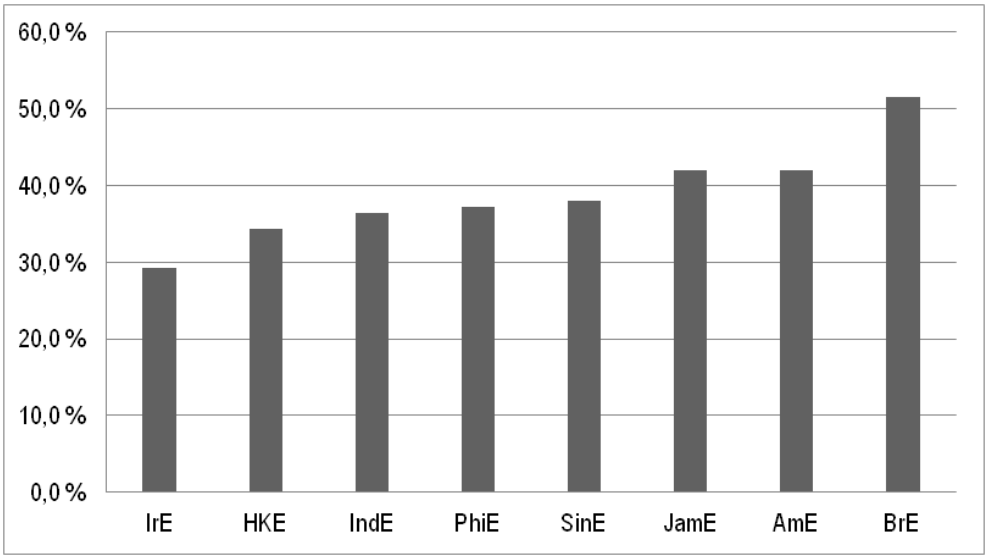

Figure 20. Prototypical progressives in World Englishes.

Looking at individual varieties in Figure 20 above, we see that BrE has the highest amount of prototypical progressives (52%), with JamE and AmE close behind (42% each).[145] The other four OC varieties are grouped together with frequencies varying from 34% in HKE to 38% in SinE. The most interesting variety here is IrE

[144] Note that, in the present study, the prototypical function includes dynamic situations only: stative and habitual situations are grouped separately. Kranich (2010), for instance, groups these three uses together under the heading 'Progressive aspect', as they, when used in the progressive, adopt a dynamic reading. As regards the present data, results comparable to Kranich's findings are presented in section 8.6 below.

[145] Raw frequencies for the distribution of all the functions are given in Table 53 in section 8.6.

which behaves very differently from the other IC varieties: the proportion of prototypical progressives is only 29%. Thus IrE uses the progressive in the less prototypical functions more often than any other variety in the present study. As discussed in section 7.1 above, the progressive in IrE is characterized as having a wider range of uses, which is reflected in the low amount of prototypical uses of the progressive.

Sentences (8.4) to (8.7) below illustrate different uses of the prototypical progressive. The most typical progressive refers to action ongoing and incomplete at the moment of speaking (as in (8.4)). However, it is not strictly necessary that the action is actually going on at the moment of speaking. In (8.5), for instance, it is possible that the speaker is not reading the book at the exact moment he or she produces the utterance, but he or she has started reading the book at one time and will continue to read the book at another time.[146] Sentences (8.6) and (8.7), on the other hand, illustrate the time-frame function of the progressive: a situation portrayed by the progressive sets up a temporal frame for another situation, which takes place during the situation portrayed by the progressive.

(8.4) Hey what you'*re doing* there Darl sit here lah (ICE-SIN, S1A-007)

(8.5) I 've read even The Hobbit and I *'m reading* Treasure Island (ICE-GB, S1A-013)

(8.6) Try them while you *'re packing* upstairs and see what how one feels (ICE-GB, S1A-022)

(8.7) When you *are dancing* you think that it is really wonderful (ICE-HK, S1A-005)

The compatibility of situation types with the progressive was discussed in sections 2.1 and 2.3 above. Used with Accomplishments, the progressive focuses the attention on the activity in progress before the goal is reached (see (8.8) and (8.9) below). Achievements, i.e. punctual and telic events, behave similarly to Accomplishments when used with the progressive: the period leading up to the change of state is put to center-stage, and whether the goal is ever reached actually remains irrelevant (see (8.10). The progressive combined with Semelfactives, punctual atelic events, results in the action being interpreted as a series of repeated acts (as in (8.11) below). In World Englishes, Activities are clearly the most common situation type to occur in the progressive (c. 60%), Accomplishments and States are fairly frequent (18% and 12%, respectively), while Achievements and

[146] Indeed, the transcript of the discussion leads us to assume that this is the case.

Semelfactives are infrequently used in the progressive (9% and 2%, respectively).[147] Thus it seems that, on average, World Englishes follow the semantic restrictions of the progressive: punctual events rarely occur in the progressive, whereas situation types that inherently refer to durative events are frequent. The fairly frequent use of States with the progressive can be regarded as evidence that the progressive is being used innovatively. Stative uses of the progressive are discussed in more detail in section 8.3 below.

(8.8) it *'s* now *getting* more and more easier because the places *are increasing* (ICE-HK, S1A-030)

(8.9) the the polarised uhm situation that you have a dance world that *is becoming* aware of the scope of dance (ICE-GB, S1A-001)

(8.10) She'*s* already *finishing* up you know (ICE-SIN, S1A-015)

(8.11) And they *were banging* their .. their soccer ball up against our.. below the bedroom .. th- the study window? (SBCSAE, sbc0002)

In the present data, the most prototypical cases of the progressive (cf. sentences (8.4) and (8.5) above) account for a great majority of progressives assigned to the category of prototypical progressives (c. 80%). The time-frame use seems to be more common in IC varieties (IC 8% : OC 3%), whereas the processive reading (i.e. of Accomplishments and Achievements used with prototypical progressive meaning) is more common in OC varieties (IC 10% : OC 13%). Semelfactives are rare in the progressive throughout World Englishes: progressives referring to a series of repeated acts (as in sentence (8.11) above) are found in only 2% of all instances of prototypical progressives.

Prototypical progressives are, as expected, the most frequent function of the progressive in World Englishes. The following subsections discuss the other functions, in the order of frequency.

[147] The proportion of situation types co-occurring with the progressive is based on a separate analysis of approximately 3,000 progressives from six ICE components. The data includes all types of progressives, regardless of the function that the tokens convey.

8.2 Subjective uses of the progressive

The progressive is increasingly being used to convey subjective rather than aspectual meaning (see e.g. Kranich 2010, Mair 2006, Smith 2005), and in particular the interpretative progressive seems to be of importance as regards the spread of the progressive in general. In addition to interpreting what is being said (as in (8.12) below), subjective progressives portray the speaker's attitude toward a situation (as in (8.13) and (8.14)), tentativeness (as in (8.15)), vividness (as in (8.16)), or it may be used as a narrative device (as in (8.17)).

(8.12) <$D> No housework to do

<$C> Mm

<$D> No hoov no floors to hoover

<$C> True Now you *'re persuading* me actually (ICE-IRL, S1A-014)

(8.13) But they *were* always in class *always whispering* to each other (ICE-JA, S1A-013)

(8.14) …cos he '*s* not fucking *showing* anywhere (ICE-IRL, S1A-019)

(8.15) And I *was wondering* if there were any opportunity to do anything like that abroad (ICE-GB, S1A-035)

(8.16) cause their .. their hips *are beating* up against you... you know (SBCSAE, sbc0002)

(8.17) And then a week later on or a couple of weeks later on she *'s walking* down the street sees his car knocks on a window (ICE-PHI, S1A-015)

In the present data, subjective progressives are the largest group of progressives that have a meaning other than the prototypical one referring to ongoing and unfinished situations: as much as 17.5% of all progressives in the data were analysed as carrying subjective meaning. As regards IC and OC varieties, subjective progressives are more frequent in the former (IC 20% : OC 16%, $p < 0.05$). Looking at individual varieties (see Figure 21 below), the varieties are clearly divided into two groups.[148] IrE, PhiE, JamE and AmE clearly have a higher proportion of subjective progressives (more than 20% each), whereas the other

[148] The position of BrE and, on the other hand, AmE on the scale presented in Figure 21 tempts one to consider the possibility that BrE usage of subjective progressives is reflected in IndE, SinE and HKE, while PhiE and JamE may have been affected by AmE.

three OC varieties and BrE use progressives to convey subjective meanings less frequently (ranging from 8% to 14%).

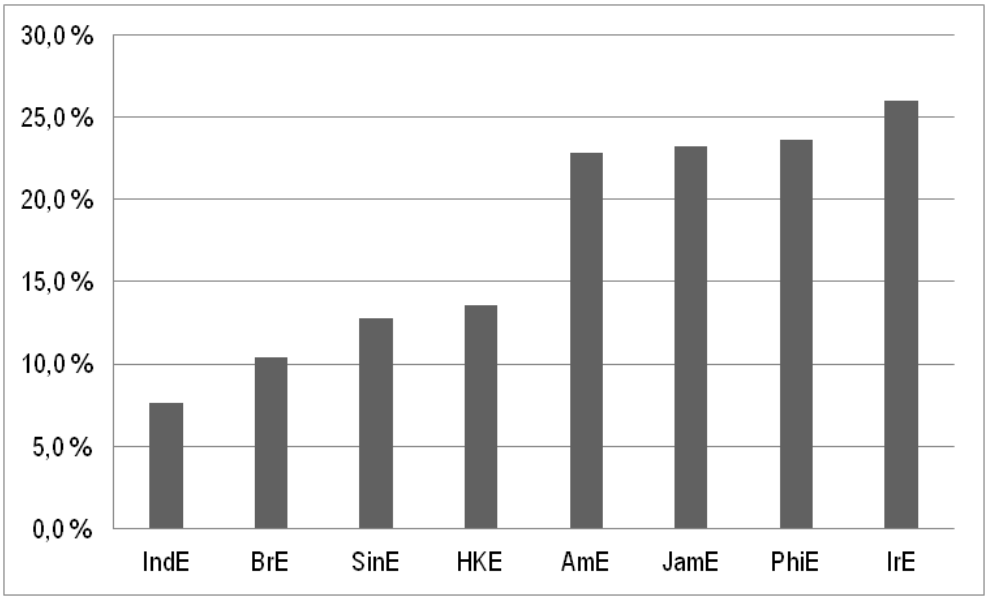

Figure 21. Subjective uses of the progressive in World Englishes.

The 350 subjective progressives found in the present data were analysed into four subcategories, as shown in Figure 22 below. In the eight World Englishes investigated, narrative progressives (including the vividness use) are the largest group with 55% of all subjective progressives on average, while attitude or emotion is conveyed in 22% of the instances. Interpretative progressives are found in 19% of all progressives carrying subjective meaning. Finally, a small proportion of subjective progressives convey tentativeness (4%). As regards the individual varieties, the use of subjective progressives attests a great deal of variation. Narrative progressives, for instance, are the largest group of subjective progressives in most varieties, excluding however BrE, in which narrative progressives are the smallest group and the interpretative the largest. In JamE and IndE, on the other hand, progressives conveying attitude or emotion are the most frequent subcategory of subjective progressives. The four subcategories are discussed in more detail in the following.

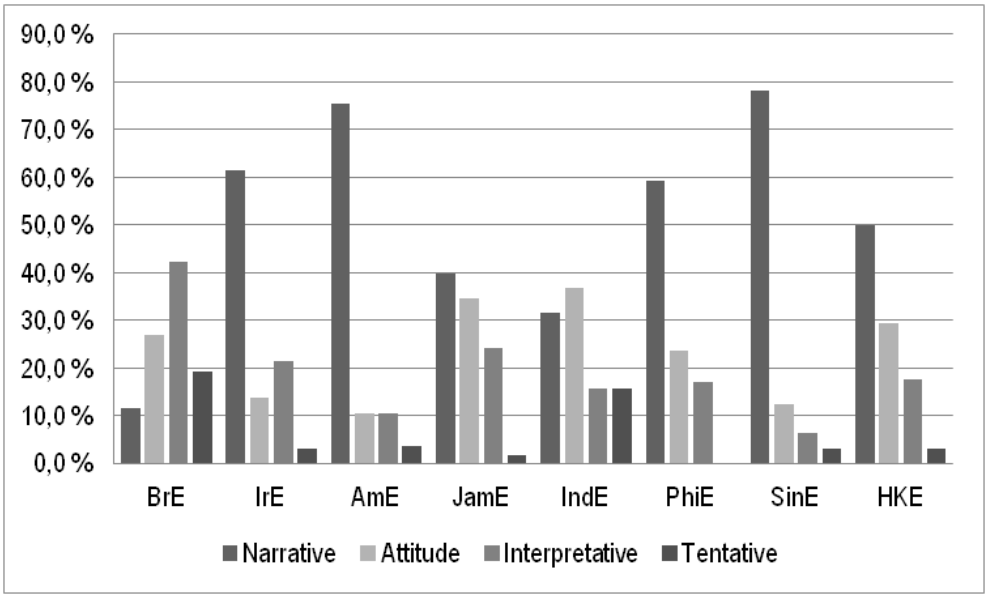

Figure 22. Subcategories of subjective progressives.

Based on the present data, the most important subcategory of subjective progressives is the narrative use of the progressive (55% of all subjective progressives). Progressives can be used to stretch the time-span of an event in order to put focus on it (as in (8.13) above, and (8.18) and (8.19) below). Progressives expressing vividness, as in (8.16) above, were included in the narrative subcategory, as they function similarly to narrative progressives.

(8.18) I tripped and fell and I looked up and the Dublin Express *was heading* straight for me (ICE-IRL, S1A-001)

(8.19) And then I '*m* just *walking* around the airport and *waiting* for somebody to pick me up (ICE-HKE, S1A-074)

Narrative progressives are found in all eight varieties investigated, but a great deal of variation exists between the varieties. BrE makes very little use of narrative progressives (11.5% of all subjective progressives in BrE), while AmE and SinE use them very frequently (75% and 78%, respectively). Narrative progressives are slightly more frequent in IC varieties (IC 58% : OC 53% of all subjective progressives, the distribution is statistically non-significant at $p = 0.2955$). It must be noted, however, that due to the relatively small data set, the conversation topics are of great importance; if the conversation recorded includes a participant telling a story of any kind, the number of narrative progressives is likely to be higher than in

a conversation which does not include storytelling. For instance, in the present data, SBCSAE, ICE-SIN and ICE-PHI were noted down as including storytelling.

A large proportion of progressives belonging to the subcategory of narrative progressives have as their main verb either *saying, talking, telling* or *thinking*. These so-called 'recentness' progressives are characterized by the co-occurrence of the past progressive and typically a temporal adverbial such as *just* (Pfaff, Bergs & Hoffmann 2013: 221). The function of recentness progressives is to put focus on the reported message, rather than on the act of reporting (ibid.). Consider (8.20) and (8.21) below:

(8.20) she *was saying* that she wants to go to Open University (ICE-SIN, S1A-005)

(8.21) you know cos I remember this morning Eddie *was telling* me that uhm when the Canadian students leave university they have choice (ICE-JA, S1A-016)

According to Pfaff et al. (ibid.: 223-4), recentness progressives are infrequent but the frequency is increasing. In the present study, recentness progressives were not included in the analysis due to their infrequency and the fact that it may be difficult to separate recentness progressives from narrative progressives.

Sometimes the only difference between the use of the non-progressive and the progressive is that the progressive conveys the speaker's attitude toward the event or act described. For instance, in sentence (8.22) below, the progressive implies the speaker's determination about the fact that he refuses to dance with other men. The non-progressive (*I don't dance with guys*) would be more neutral and lack the expression of attitude. In (8.23), the progressive emphasizes the speaker's desperation and is thus emotive in function.

(8.22) I'*m* not *dancing* with guys. (SBCSAE, sbc0002)

(8.23) I *'m drowning* with Accounting numbers with Accounting teachers (ICE-PHI, S1A-013)

This subcategory also includes subjective progressives with ALWAYS, used to express negative speaker attitude. As Kranich (2007) points out, it is also possible that the construction expresses positive or neutral attitude. In the present data, there are only eight instances of the progressive co-occurring with the adverbials *always* (six instances), *constantly* (once) or *all the time* (once): four of these convey negative attitude toward what is said (cf. (8.13) reproduced below), while the other four are more neutral in character (as in (8.24) below). The eight occurrences of

subjective progressives with ALWAYS are all found in the OC varieties (three instances in HKE, two in JamE, and one in IndE, PhiE and SinE each).[149]

(8.13) But they *were* always in class *always whispering* to each other (ICE-JA, S1A-013)

(8.24) And I think the law system *is always changing* (ICE-HK, S1A-013)

Progressives expressing attitude and emotion in general are the second largest subgroup of subjective progressives (22%), with proportions ranging from 10.5% in AmE to 35% in JamE and 37% in IndE. Attitudinal progressives are more frequent in the OC varieties (IC 15% : OC 27%, $\chi^2(1) = 7.3079$, $p < 0.01$).

The extract in (8.25) below exemplifies the function of interpretative progressives: speaker C uses the progressive *I'm saying* to ensure that the other participants understand what she wants to convey – in other words speaker C *interprets* her first utterance.

(8.25) <$C> It's very nice you know you just lie down there and then you can wear sort of like bikini you know

<$B> Ah really

<$A> My goodness

Wah you become so open when you are there

<$C> No lah I mean I don't have the figure for bikini

I mean I'*m saying* you can

<$B> Because she won't she won't meet anybody else except us

<$C> No

I'*m saying* you can if you want to

(ICE-SIN, S1A-011)

Interpretative progressives are found in all eight varieties, and are clearly most frequent in BrE (42% of all subjective progressives in BrE). JamE and IrE also use progressives to interpret what was said (24% and 22%, respectively), whereas SinE and AmE make use of these progressives infrequently (6% and 10%, respectively). The proportion of interpretative progressives of all subjective progressives is 21%

[149] However, subjective progressives with ALWAYS are found in IC varieties, as well, when the data is considered as a whole, i.e. 100,000 words per corpus.

in IC varieties, and 17% in OC varieties (statistically non-significant at $p = 0.3925$). The present results are interesting regarding the argument made by Kranich (2010), among others, concerning the role of interpretative progressives in the overall spread of the progressive. Indeed, it seems that interpretative progressives may play a role, but only in BrE in which the proportion of interpretative progressives of all subjective progressives is close to 50%. However, when the number of interpretative progressives is related to the number of all progressives, it quickly becomes clear that the present data does not support Kranich's argument: the proportions are extremely low in all varieties (0.3% to 1.6%). Most likely a larger set of data is needed to verify the role of interpretative progressives in the increasing frequency of the progressive overall.

Finally, a small amount of subjective progressives in the present data are used to render a request more polite, as in (8.26), or to render an utterance more tentative, as in (8.27) below.

(8.26) I just *was wondering* whether we could discuss our projects so we know what we are all doing (ICE-SIN, S1A-033)

(8.27) I *'m hoping* to go at Easter (ICE-IRL, S1A-016)

BrE and IndE, in particular, use the progressive to convey tentativeness: as much as 19% of all subjective progressives in BrE and 16% in IndE are found in this subcategory. However, it must be noted that the raw frequencies of tentative progressives are extremely low (15 in total), so that it is impossible to make any generalizations on the use of tentative progressives in general. Tentative progressives were assigned into their own subcategory, as a number of previous studies (e.g. Römer 2005) discuss their use. However, they could be added to the subcategory of attitude, as tentativeness can be thought of as portraying an attitude towards what is being said.

In sum, the largest group of progressives referring to a situation other than ongoing or incomplete is formed by those conveying subjective meaning – narrative, attitude, interpretative or tentative. There is a clear distinction in the frequency of subjective progressives between IC and OC varieties: the function is more frequent in IrE and AmE than in the OC varieties (but infrequent BrE). The present results do not seem to support the argument made by Kranich (2010) and Smith (2005), among others, according to which the interpretative progressive could be affecting the overall frequency of the progressive. Nevertheless, the fact that subjective progressives are the second most frequent group of progressives in

the World Englishes included in the present data indicates that they are a legitimate bundle of functions separate from the aspectual uses of the progressive.

8.3 Stative uses of the progressive

As discussed in section 2.3 above, when stative verbs combine with the progressive, the resulting meaning is, for instance, that of temporariness (as in (8.28) below) or deliberate action (as in (8.29) below).

(8.28) I think the place I *'m living* now is tut quite strange for me (ICE-HK, S1A-026)

(8.29) *Are* you *being* funny (ICE-IRL, S1A-020)

The stative verbs found in the present data were categorized into semantic classes (see section 2.3 above); approximately 40% represent verbs of cognition, emotion or attitude (see (8.30) below), 35% are stance verbs (as in (8.28) above), 19% represent verbs of being and having (as in (8.29) above), and less than 7% are verbs of perception and sensation (see (8.31) below). The proportions of the different types of stative verbs are rather similar to the results obtained by Leech et al. (2009), who find only a slight increase in the amount of stative progressives between the 1960s and the 1990s. A number of stative progressives (mainly verbs of cognition, emotion or attitude) in the present data were labelled as subjective progressives because they express tentativeness (see section 8.2 above).[150]

(8.30) I *was* half sort of *expecting* it was going to be like uh a murder picture (ICE-IRL, S1A-007)

(8.31) So I mean it *was looking* promising becau but they couldn't go forward with it (ICE-JA, S1A-010)

[150] Kranich (2010) avoids the problem of overlapping functions by discussing the use of stative verbs with the progressive separately from the functions of the progressive and integrating stative progressives into the 'Progressive aspect' category. To a great extent, the analysis of stative progressives is a *lexical* categorization, as the analysis is mostly based on locating stative main verbs (as well as dynamic main verbs used statively). The category of stative progressives thus differs to a degree from the other categories, as it is the main verb rather than the function which determines the inclusion of a progressive token into the category.

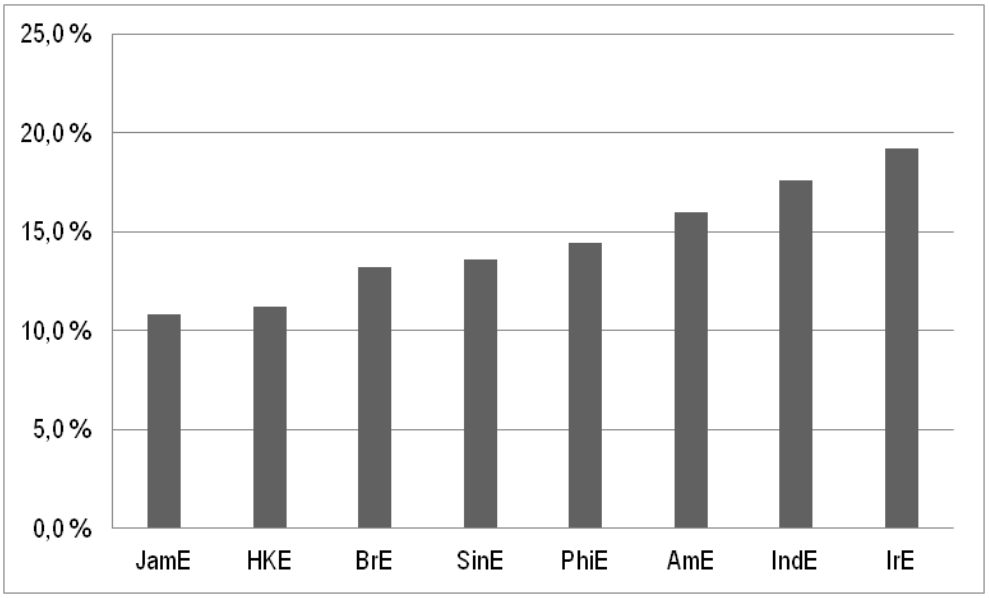

Figure 23. Stative uses of the progressive in World Englishes.

Figure 23 above represents the frequency of stative progressives in the individual varieties; stative progressives are most frequent in IrE (19% of all progressives in IrE) and IndE (18%) and least frequent in JamE and HKE (11% each). This result is somewhat surprising, as according to Kortmann and Lunkenheimer (2013), among others, one would have expected to see more OC varieties at the far right-hand end of Figure 23. The present results thus seem to lend only partial support to the notion that OC Englishes are characterized by frequent use of stative verbs in the progressive. IrE and AmE have clearly higher frequencies of stative progressives than most OC varieties included in the present study. However, stative progressives have two uses, one of which is considered standard and the other non-standard according to whether they refer to situations that are temporally delimited (as in (8.32) below) or temporally non-delimited (as in (8.33) below).

(8.32) he *was standing* very close behind us you know (ICE-SIN, S1A-031)

(8.33) But they *must* be uh *be belonging* to some I mean uh good I mean families (ICE-IND, S1A-025)

The proportions of standard and extended stative progressives in the present data are shown in Figure 24 below. Most varieties seem to follow the prescriptive rules

of how stative verbs should be combined with the progressive, except for IndE and HKE, as well as JamE to a lower extent, which show more variation and use stative progressives in extended contexts. The proportion of extended stative progressives is approximately the same in IndE and HKE (36% and 32%, respectively) – these two varieties clearly stand out as allowing the progressive to be used in situations that are stative and temporally non-delimited, i.e. opposite to the semantic restrictions related to the prototypical progressive. This result thus forces one to re-assess the statement made above regarding stative progressives in OC varieties; it is the use of *extended* stative progressives that sets OC varieties (IndE and HKE in particular) apart from IC varieties ($\chi^2(1) = 10.3866$, $p = 0.001$). Of all stative progressives in IC varieties, only 8% are extended, while the corresponding figure for OC varieties is 21%.[151]

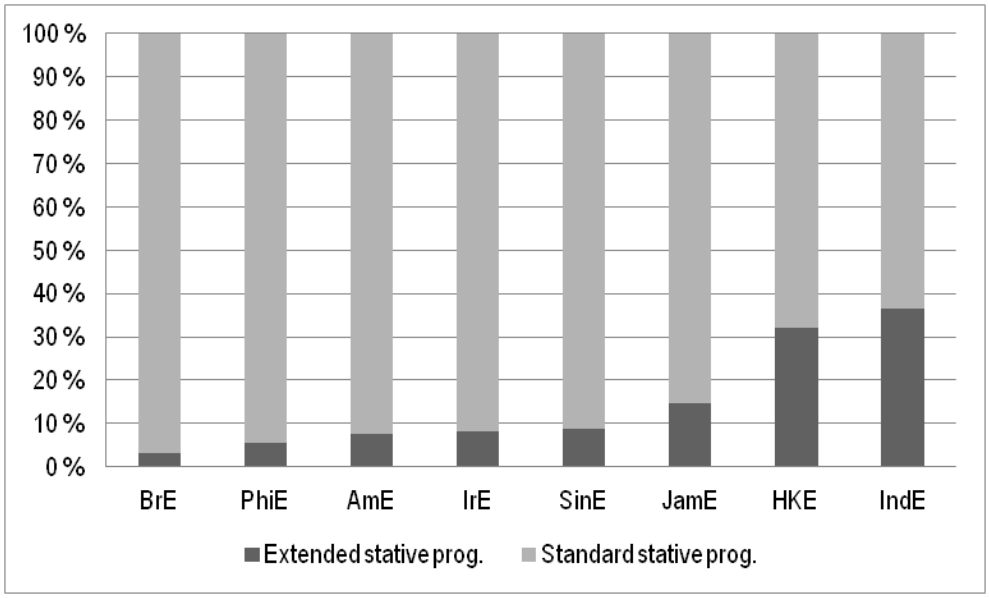

Figure 24. Extended and standard stative uses of the progressive in World Englishes.

Sentences (8.34) and (8.35) below illustrate the use of extended stative progressives. In (8.34), the progressive is used to refer to a permanent situation: it

[151] Of *all* progressives (all functions), 1% in IC varieties and 3% in OC varieties fall into the category of extended stative progressives.

is not likely that waterfalls are a temporary phenomenon.[152] As regards (8.35), the verb STAY can be used to refer to permanent residence in SinE (Deterding 2007: 81), and apparently also in IndE. The speaker is introducing herself and tells the other participants that she lives in Ambala.

(8.34) You have you *must be having* waterfalls? (ICE-IND, S1A-008)

(8.35) I'm from Haryana I'*m staying* in Ambala My mother tongue is Hindi (ICE-IND, S1A-007)

According to Kortmann and Lunkenheimer (2013), extended use of stative verbs is "pervasive or obligatory" in IndE, but "extremely rare" in HKE and colloquial SinE. The other varieties included in this study are neutral as regards extended stative progressives in Kortmann and Lunkenheimer's report. The results of this study confirm that the use of extended stative progressives in IndE is indeed "pervasive or obligatory", but the informants' view and the present results concerning extended stative progressives in HKE differ to a great extent. The finding of the present study is supported by data from ICLE-CH[153], which also indicates high frequency of extended stative use in HKE (see Meriläinen et al.: fc.). The matter is further discussed in section 8.6 below.

8.4 Futurate uses of the progressive

Futurate use of the progressive without *will* or *shall* (cf. (8.36) below)) may be thought of as "a metonymic extension of the basic meaning of the progressive" (Leech et al. 2009: 133), as the situation referred to is often already being planned and regarded as already in progress or as just about to start.

(8.36) When *'s* your Mum *coming* back (ICE-GB, S1A-006)

As Mair and Hundt (1995), and Leech et al. (2009) point out, it is sometimes difficult to tell whether the situation referred to has already begun (and is not futurate anymore) or is just about to begin (and therefore futurate). In the present

[152] The discussion in file S1A-007 (ICE-IND) concerns travelling and beautiful places to visit, rather than temporary conditions related, for instance, to exceptional flooding.

[153] The Chinese component of the *International Corpus of Learner English* can be regarded as representative of HKE, as well.

study, it was on most occasions possible to determine if the progressive in question was futurate or not – otherwise the progressive was categorized as 'Indeterminate'. Consider, for instance, sentence (8.37) below. In isolation, it is difficult to tell whether the speaker is already on his or her way, or if he or she is planning the action. However, the context reveals that the utterance is an answer to a question *What are you gonna do after here?* and is thus clearly futurate.

(8.37) Well I *'m going* to our store no first to the church (ICE-PHI, S1A-013)

Another type of the futurate progressive is the combination of *will* or *shall* with the progressive (see (8.38) below), which indicates that, in the future, the situation referred to will be in progress, regardless of the intention of anyone concerned (Kranich 2010: 182). Describing a situation as being in progress in the future is, according to Kranich (ibid.), more definite than saying that the situation will occur. Considering sentence (8.38), *doing my proposal* is described as a situation already in progress next year; the speaker perhaps wants to ascertain the hearer that he or she is will definitely write the research proposal (*hopefully* somewhat softens the determination).

(8.38) And then I come back next year hopefully I'*ll be doing* my proposal for my research (ICE-JA, S1A-020)

The historical development of futurate progressives is that of increasing frequency (Nesselhauf 2007; Kranich 2010). Kranich's (2010) data indicates that the use of *will/shall* + progressive with future reference was nearly non-existent until the 20th century, while the use of progressives indicating futurity without modals started to increase slightly earlier, at the mid-19th century (p. 186, Table 23). Mair and Hundt (1995) and Nesselhauf (2007) both propose that the increasing use of futurate progressives may have affected the general increase of the progressive form; Leech et al. (2009: 133) point out that the impact is probably small, as the number of futurate progressives is low. Kranich (2010: 186) lends support to both suggestions as her data shows a clear increase in the proportion of futurate progressives, although the frequency overall remains low.

Figure 25 below shows the proportion of futurate uses of all progressives in the present data. SinE, IrE and IndE have the highest frequencies of futurate progressives (19%, 17% and 16%, respectively), while the lowest frequencies are found in AmE and JamE (5% and 6%, respectively). With regard to futurate progressives in IC and OC varieties, there is no statistically significant difference between the two groups: the proportion of futurate progressives in IC varieties is

approximately 10%, and the corresponding figure for OC varieties is 12% (p = 0.1464). Looking at the main verbs of futurate progressive verb phrases, it quickly becomes evident that the verbs *coming, going* and *leaving* account for nearly 60% of all futurate progressives. *Going* is used in the futurate in all varieties, but *coming* is almost exclusively found in SinE, and *leaving* in IndE. Although the use of futurate progressives seems to be lexically constricted, the slight difference between IC and OC varieties remains valid – futurate progressives are more frequently found in OC varieties even if *going, coming* and *leaving* are excluded from the analysis.

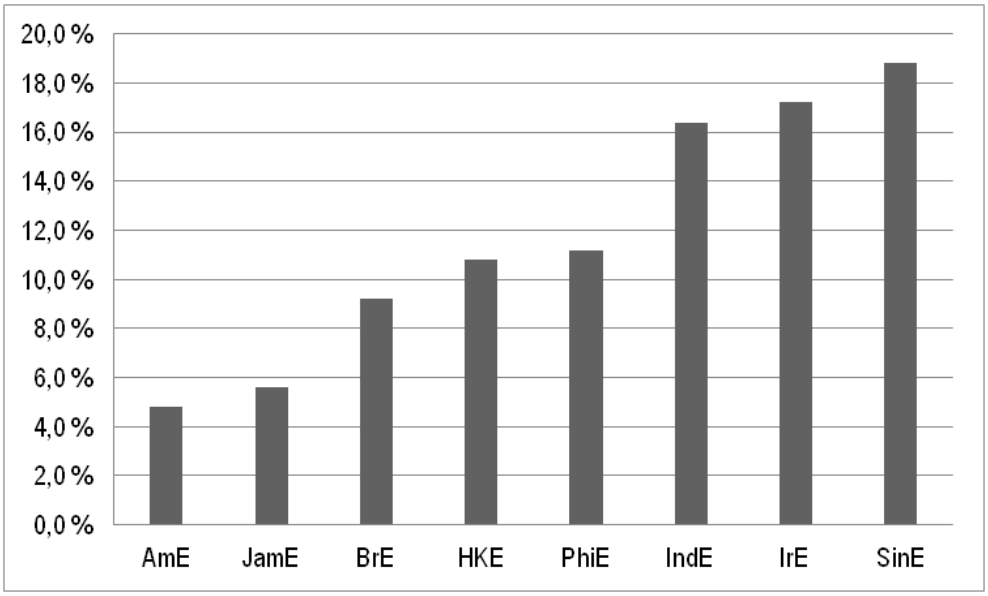

Figure 25. Futurate uses of the progressive in World Englishes.

Most of the futurate progressives (c. 76%) in the data refer to future events without modal auxiliaries *will* or *shall*, which conforms to Nesselhauf and Römer's (2007) findings. Thus the main function of futurate progressives seems to be 'expression of near future' (as in (8.39) and (8.40) below), while the 'future as a matter of course' meaning, expressed by *will/shall* + *be* V*ing*, remains marginal (c. 24%) (see (8.41) below). Figure 26 below represents the proportions of the two types of futurate progressives in individual varieties. In most varieties, the distribution of the two types is stable (ranging from 23% in SinE to 30% in HKE), but BrE and JamE clearly have smaller shares (4% and 14%, respectively) of the second type of futurate progressives (with a modal).

(8.39) Yeah in any case I *am leaving* at two twenty in the afternoon (ICE-IND, S1A-019)

(8.40) They *are flying* in three models from Milan (ICE-HK, S1A-014)

(8.41) Ah will you have a wee tear in your eye Well I *won't be seeing* her going (ICE-IRL, S1A-001)

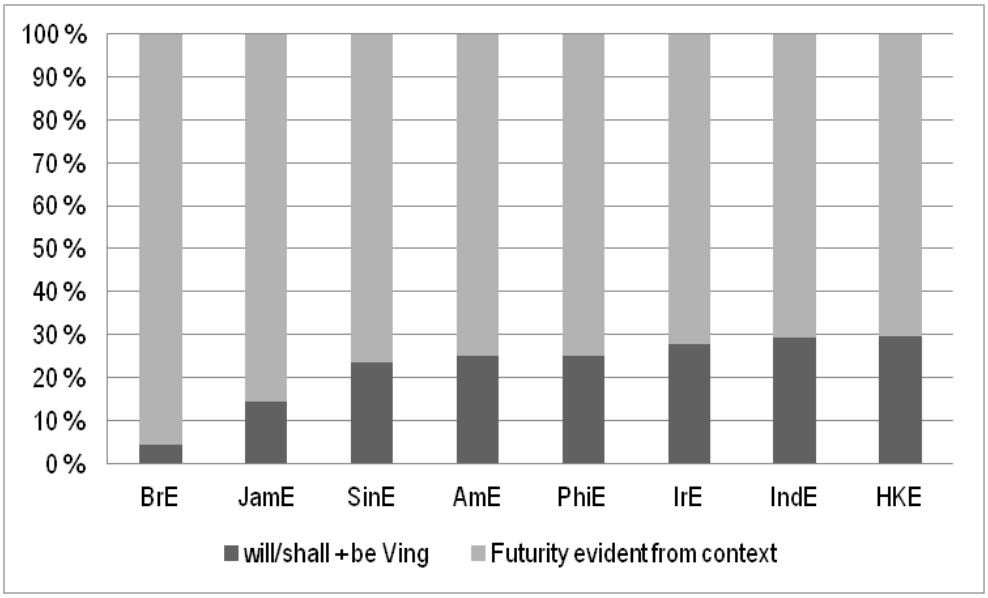

Figure 26. Subcategories of futurate progressives.

When tagging the data for purposes of the present study (see section 4.4 above), all future expressions were marked separately, i.e. it is possible to investigate how World Englishes refer to future situations in general. As Figure 27 below shows, the most common way of referring to future situations is by using *will* or *shall* and the infinitive (59%), while the *be going to* with the infinitive is also a frequently used option.[154] Progressives with future reference (without *will* or *shall*) account for approximately 9%, while the present simple is found in 6% of all futurate instances. Finally, the second type of futurate progressives is the least frequent (3%) expression of futurity in general. Thus in total, future events are referred to with the progressive in approximately 12% of all future expressions, which is rather a small proportion.

[154] The high frequency of *be going to* + infinitive confirms that the decision to exclude the construction from the analysis as a whole was correct. See section 4.3 above.

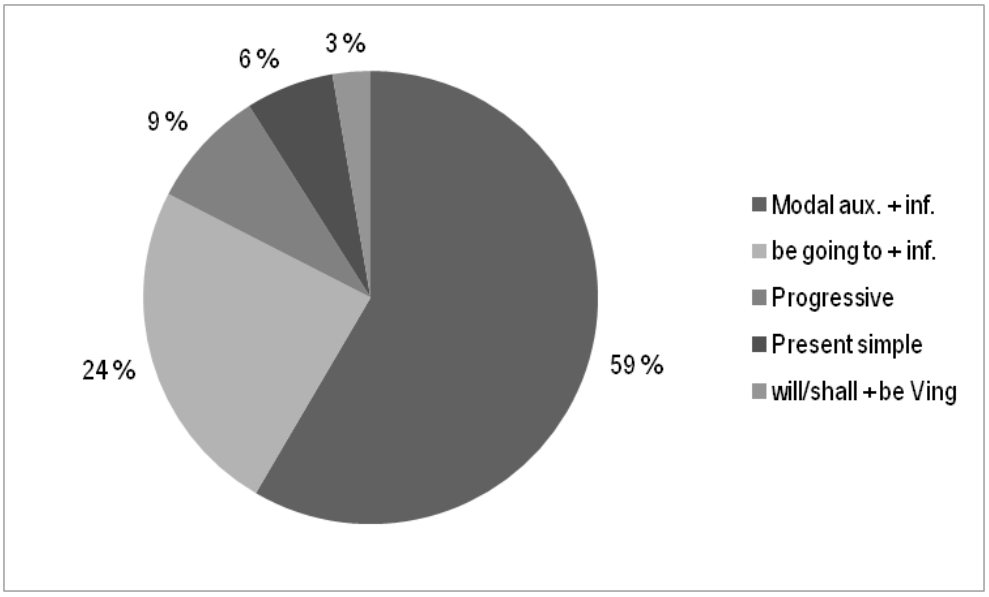

Figure 27. Future expressions in World Englishes.

A diachronic view of how the use of futurate progresssives has developed during the recent decades in World Englishes would perhaps reveal whether futurate progressives can indeed be held accountable for the general increase in the frequency of the progressive, as suggested by Mair and Hundt (1995) among others. This is, however, a matter that cannot be pursued in the scope of this study, but surely warrants further research.

8.5 Habitual uses of the progressive

Of the 2,000 progressives under scrutiny 8.2% were categorised as habitual progressives, i.e. as consisting of series of events (Leech and Svartvik 1994: 66). Habitual progressives in the present data refer to habits that are temporary (as in (8.42) below) or of more permanent kind (as in (8.43) below). The two types of habitual progressives are discussed in more detail towards the end of this section.

(8.42) I *was shopping* for about three days (ICE-SIN, S1A-003)

(8.43) Like my family all the people they *are taking* non-veg (ICE-IND, S1A-007)

The proportion of habitual progressives is clearly smaller than that of any other function related to the progressive, which may be seen as indication of habitual progressives not having crystallized as a legitimate possibility in the speakers' minds. As becomes evident in Gachelin's (1997) article, there are a number of other means of expressing habituality in standard and non-standard varieties of English, which may, in part, explain the relatively low frequency of habitual progressives in the present data. Also Kranich (2010: 172-175) finds that habitual progressives remain infrequent throughout the past four centuries, although a slight rising tendency is discernible.

There is no significant difference between IC and OC usage of habitual progressives: the proportion of habitual progressives is approximately 8% for both groups of varieties ($p = 0.7198$). However, a look at the individual varieties does reveal some differences: habitual progressives occur clearly most frequently in AmE (c. 11%), and the percentage of habitual progressives in most varieties varies between 6% and 10% (see Figure 28 below). Quite surprisingly, IrE has the lowest amount of habitual progressives (c. 4%). Celtic Englishes in general are considered to make extensive use of habitual progressives, even to such extent that Gachelin (1997: 40) refers to a grammar book of Welsh English which gives BE + V*ing* as a marker of "habitual present tense". Similarly, Meriläinen et al. (fc.) find that habitual progressives are frequent in Welsh English and in traditional dialects of BrE represented in SED.[155] A possible explanation for the low frequency of habitual progressives in IrE is the fact that IrE (and other Celtic Englishes) uses other constructions to express habituality (e.g. *He bees waiting, He does be waiting*, see Gachelin (1997) for more information).[156]

[155] Survey of English Dialects, more information at http://sounds.bl.uk/accents-and-dialects/survey-of-english-dialects.

[156] There are no instances of *bees* in the spoken component of ICE-IRL (600,000 words), and one instance of *does be* (S1A-087).

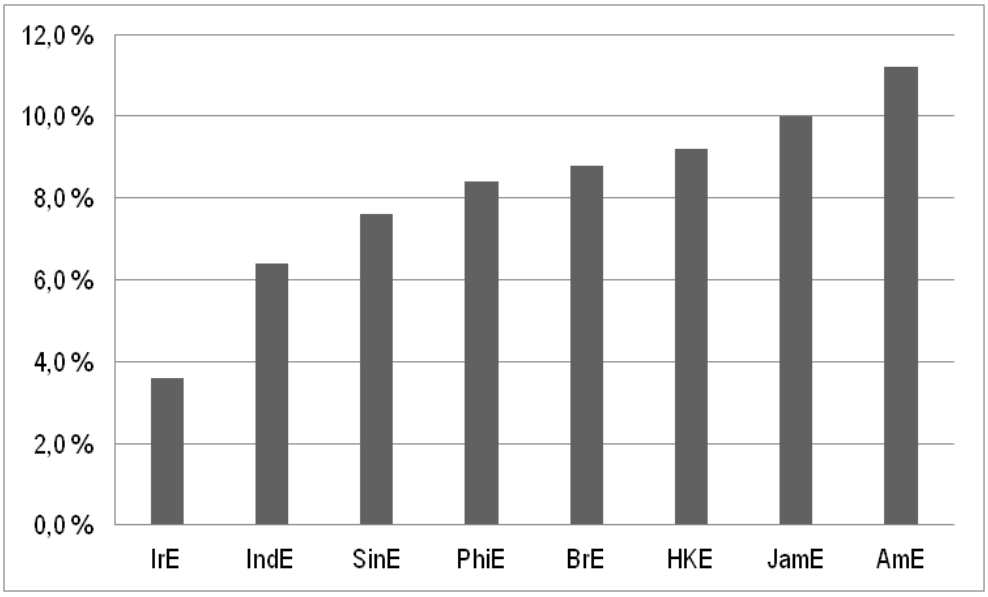

Figure 28. Habitual uses of the progressive in World Englishes.

The standard use of habitual progressives refers to habits that are limited in time, i.e. temporary (see section 2.3 above). The temporariness of the habit may be indicated by a temporal adverbial or temporal clause (as in (8.44) below) or it may be evident from the context (as in (8.45) below).

(8.44) He would he *was* just *calling* me up for barely two months (ICE-PHI, S1A-019)

(8.45) And uhm and at the College of Speech Sciences I *'m putting* things in alphabetical order for them (ICE-GB, S1A-011)

Extended habitual progressives, in contrast, refer to situations that are not limited to a specific period of time but are rather of 'general currency' (consider (8.46) and (8.47) below).

(8.46) what you say like a countryside the people are uh uh *are riding* horses they they haven't much cars (ICE-HK, S1A-042)

(8.47) Your snake *was eating* goldfish? (SBCSAE, sbc0015)

Figure 29 below shows the proportions of standard habitual progressives and extended habitual progressives in the eight varieties. On average, 13.5% of all habitual progressives in the data refer to unlimited habitual situations, but there is considerable variation between the varieties. IrE does not have any instances of

extended habitual progressives, and extended instances are infrequent in AmE and BrE, and SinE and PhiE (ranging from 4% to 10%). JamE and HKE, which use habitual progressives in general frequently, include a relatively high number of extended habitual progressives, as well (20% and 22%, respectively). In IndE, as much as 44% of all habitual progressives are of the extended kind. The most interesting cases are inarguably IndE and AmE. AmE has the highest amount of habitual progressives in general but only one of the 28 instances is extended, while the use of habitual progressives in IndE is clearly non-standard (7/16 extended). This result conforms to Kortmann and Lunkenheimer (2013): extension of the progressive to habitual situations is pervasive or obligatory in IndE and absent in (colloquial) AmE. Another interesting finding is the fact that extended use of habitual progressives and, on the other hand, of stative progressives are both frequent in IndE and HKE. Thus these two varieties seem to be clearly bending the semantic restrictions of the progressive and using it in truly extended contexts.

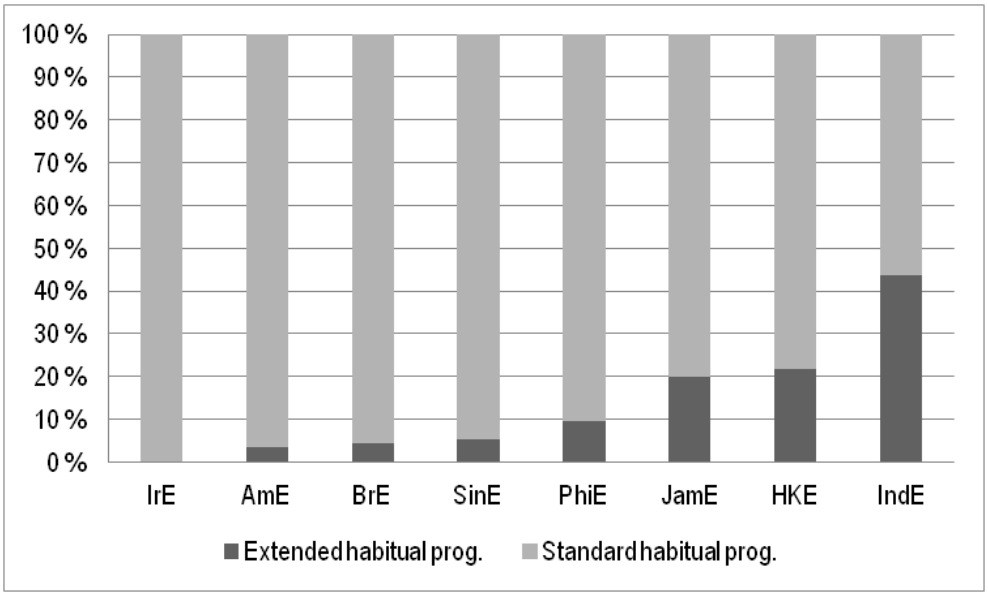

Figure 29. Extended and standard habitual uses of the progressive.

While standard habitual progressives are more frequently found in IC varieties in the present data, extended habitual progressives are significantly more frequent in OC varieties (IC 3% of all habitual progressives : OC 19%, $\chi^2(1) = 8.0908$, $p <$

0.01).[157] IC varieties can thus be said to follow the traditional rules of not combining the progressive with non-delimited habitual situations, while OC varieties seem to be disregarding the rule and using the progressive innovatively in situations where the semantic restrictions of the form need to be extended. The present results only partly conform to those presented by Kortmann and Lunkenheimer (2013): extended use of habitual progressives is a frequent feature in IndE, but the high frequency in HKE is, again, surprising. See section 8.6 below for further discussion.

8.6 Summary of results

Chapter 8 has discussed the functions of the progressive in World Englishes. On the whole, a rather stable picture emerges: the prototypical progressive is the most frequent function in all varieties, albeit with a small marginal to the subjective progressive in IrE (see Figure 30 below, detailed information given in Table 53 below). Subjective progressives are in most varieties in second rank, excluding BrE in which the second rank goes to stative progressives, and IndE and SinE in which both stative and futurate progressives are more frequent than subjective progressives. Futurate progressives are clearly more frequent in SinE (2nd rank) than in any other variety. Finally, habitual progressives are the least frequent category in most varieties. Based on the information presented in Figure 30, it seems that SinE differs the most from other varieties: the ranking of two categories (futurate and subjective) differs from the average.

[157] Of *all* progressives (all functions), extended habitual progressives account for 0.3% in the IC varieties and for 1.6% in the OC varieties.

Table 53. Functions of the progressive in World Englishes (N = 250 per variety).

	Progressive	Subjective	Stative	Futurate	Habitual	Other uses	Indeterminate
BrE	51.6% (129)	10.4% (26)	13.2% (33)	9.2% (23)	8.8% (22)	4.8% (12)	2.0% (5)
IrE	29.2% (73)	26.0% (65)	19.2% (48)	17.2% (43)	3.6% (9)	4.8% (12)	0
AmE	42.0% (105)	22.8% (57)	16.0% (40)	4.8% (12)	11.2% (28)	2.0% (5)	1.2% (3)
JamE	42.0% (105)	23.2% (58)	10.8% (27)	5.6% (14)	10.0% (25)	3.6% (9)	4.8% (12)
IndE	36.4% (91)	7.6% (19)	17.6% (44)	16.4% (41)	6.4% (16)	14.0% (35)	1.6% (4)
PhiE	37.2% (93)	23.6% (59)	14.4% (36)	11.2% (28)	8.4% (21)	4.4% (11)	0.8% (2)
SinE	38.0% (95)	12.8% (32)	13.6% (34)	18.8% (47)	7.6% (19)	5.6% (14)	3.6% (9)
HKE	34.4% (86)	13.6% (34)	11.2% (28)	10.8% (27)	9.2% (23)	19.6% (49)	1.2% (3)
TOTAL	38.9% (777)	17.5% (350)	14.5% (290)	11.8% (235)	8.2% (163)	7.4% (147)	1.9% (38)

Variety	Progressive	Subjective	Stative	Futurate	Habitual
BrE	1st rank	3rd rank	2nd rank	4th rank	5th rank
IrE	1st rank	2nd rank	3rd rank	4th rank	5th rank
AmE	1st rank	2nd rank	3rd rank	5th rank	4th rank
JamE	1st rank	2nd rank	3rd rank	5th rank	4th rank
IndE	1st rank	4th rank	2nd rank	3rd rank	5th rank
PhiE	1st rank	2nd rank	3rd rank	4th rank	5th rank
SinE	1st rank	4th rank	3rd rank	2nd rank	5th rank
HKE	1st rank	2nd rank	3rd rank	4th rank	5th rank
Total	1st rank	2nd rank	3rd rank	4th rank	5th rank

Figure 30. The order of the functions of the progressive in individual varieties, according to relative frequency.

In addition to the high frequency of subjective progressives overall and of futurate progressives in OC varieties, an interesting feature of the progressive in World Englishes is its extended use with statives and habituals. In order to better understand the extent of extended use, the present data was re-arranged into three categories: aspectual uses, extended uses and subjective uses (see Table 54 below). Aspectual uses are those that are compatible with the semantic restrictions of the progressive: prototypical progressives, futurate progressives, and the standard, or delimited, use of stative and habitual progressives. Extended, or non-delimited, use of stative and habitual progressives belongs to the category 'Extended use', while 'Subjective uses' corresponds to subjective progressives. Other and indeterminate tokens are grouped together. Table 54 below shows that aspectual uses of the progressive are in the clear majority in each variety, with the frequency ranging from 60% in HKE to 82% in BrE. Aspectual uses of the progressive are slightly favoured in IC varieties (IC 74% : OC 68%), which indicates that non-aspectual uses are more frequent in OC varieties. Subjective uses were discussed in section 8.2 above: IC varieties use progressives to convey subjective meaning more often than OC varieties (IC 20% : OC 16%). Extended use, on the other, characterizes OC varieties (IC 1% : OC 4%), or to be more precise, IndE and HKE. In these two varieties, 9% and 6%, respectively, of all progressives refer to habits or states that are not regarded as temporary but permanent, and thus generally speaking incompatible with the progressive. This is a widely attested feature of IndE (see e.g. Balasubramanian 2009, Kortmann and Lunkenheimer 2013), and can be traced

back to the fact that speakers of IndE transfer aspectual system from their mother tongues into English, thus resulting in high frequency of the progressive, but also high frequency of extended use (see section 5.3 above). However, the fact that HKE acts similarly is somewhat surprising. The Chinese substrate languages spoken by speakers of HKE have an aspectual system different from that of English, which combined with the fact that HKE speakers may not be as proficient as speakers of SinE (see e.g. Bolton 2002a), for instance, could result in HKE speakers disregarding the semantic restrictions of the English progressive form. High frequency of extended use is found also in ICLE-CH, which further indicates that speaker proficiency might be an issue in this case (see Meriläinen et al.: fc.).

Table 54. Frequency of aspectual, extended and subjective progressives in World Englishes.

Variety	Aspectual uses	Extended uses	Subjective uses	Other+Indet	Total
BrE	82.0% (205)	0.8% (2)	10.4% (26)	6.8% (17)	100% (250)
IrE	67.6% (169)	1.6% (4)	26.0% (65)	4.8% (12)	100% (250)
AmE	72.4% (181)	1.6% (4)	22.8% (57)	3.2% (8)	100% (250)
JamE	64.8% (162)	3.6% (9)	23.2% (58)	8.4% (21)	100% (250)
IndE	67.6% (169)	9.2% (23)	7.6% (19)	15.6% (39)	100% (250)
PhiE	69.6% (174)	1.6% (4)	23.6% (59)	5.2% (13)	100% (250)
SinE	76.4% (191)	1.6% (4)	12.8% (32)	9.2% (23)	100% (250)
HKE	60.0% (150)	5.6% (14)	13.6% (34)	20.8% (52)	100% (250)
Total	70.1% (1,401)	3.2% (64)	17.5% (350)	9.3% (185)	100% (2,000)

Although the results of the study do provide a view of the functions that the progressive is used in World Englishes, some methodological problems need to be taken into account when analyzing the results. First and foremost, the categorization of semantic functions is susceptible to subjective, rather than objective, decisions. This is a fact that cannot easily be avoided, as the researcher is detached both from the person who produced the progressive and from the situation it was produced in. It is thus difficult, if not impossible, to decide which function the speaker originally intended. In some instances, the decision is easier to make with the help of contextual or morphosyntactic information, as is the case with futurate progressives. In other instances, lexical cues determine the main category into which a progressive token should be placed, but even then the researcher needs to consider the temporariness or permanence of the situation (e.g.

with stative progressives). The chance of making incorrect analyses is the greatest with subjective progressives, as on many occasions, the subjective nature of a progressive would be more easily detected at the time it is uttered. Furthermore, the definitions of subjective progressives may be difficult to pinpoint, so that the analysis is further problematized by hazy subcategories. The present results are further limited by a small set of data, 250 progressives per corpus, and possibly also by the fact that the tokens come from a limited number of files. It is possible that progressives from a single file, or even from a single speaker, affect the result of the analysis to a great extent. More representative results would have perhaps been obtained by a more random sample, by including tokens from as many files as possible. Thus the results of this study, as well as of any other similar study, should be considered indicative rather than conclusive. Nevertheless, the results presented in this chapter do provide us with a view of how the functions of the progressive transpire in World Englishes.

9 Conclusion

The present study has investigated the frequency, morphosyntactic variation, lexical variation and the functions of the progressive in eight World Englishes. The number of features and varieties studied makes the study one of the most comprehensive investigations into the progressive: although the construction has been extensively studied in the past decades, a World Englishes viewpoint has so far been mostly neglected. In this concluding chapter, the results of the study are briefly summarized, and interrelated to each other in order to gain a general view over the progressive in World Englishes (section 9.1). Section 9.2 discusses the success or failure of the methodological choices, and finally, section 9.3 sets our sight on further research on the progressive.

9.1 Results on the progressive in World Englishes

The research questions set in Chapter 1 are reproduced below, and answers to them are provided in the following.

a) What is the frequency of the progressive in World Englishes: e.g. is the progressive more frequently used in Outer Circle varieties?

b) What is the grammatical environment of the progressive: e.g. are there differences in the verb phrase patterns in different varieties of English?

c) Which main verbs is the progressive used with: e.g. are the verbs shared in World Englishes?

d) What are the functions of the progressive in World Englishes: e.g. do Outer Circle Englishes extend to non-standard uses?

As regards the frequency of the progressive in World Englishes, the results of the study show that the progressive is *not* more frequently used in OC varieties than in IC varieties, although some individual OC varieties do attest more frequent use of progressives than BrE and AmE. On the whole, HKE has the lowest frequency of progressives (M=373), while IrE has the highest (M=1,073). Most varieties attest very little differences, having a frequency of M=789 on average, which further

highlights the low frequency found in HKE. Section 5.3 above discussed the role of substrate transfer as a possible reason for the differences in the frequency of the progressive: it was found that the progressive aspect may be expressed with other means in addition to BE + V*ing* in HKE, arising from the fact that the progressive marker *gan* in Cantonese is optional. Substrate transfer was found to affect the frequency of the progressive also in IndE and SinE, as well as IrE.

The verb phrase patterns in which the progressive occurs in World Englishes can be arranged into three groups: in the first group, simple progressives are used in more than 90% of all instances, in the second, complex forms of the progressive are more frequent than in the other varieties, and in the third group, fragmentary forms, alongside the complex, account for a large chunk of all instances. The way the varieties investigated are divided into these groups is interesting: the most traditional varieties, i.e. BrE, IrE and AmE, are found in Group 1, IndE and PhiE in Group 2, and finally JamE, SinE and HKE in Group 3. This grouping suggests that the progressive may be increasing in IndE and PhiE through the complex forms, i.e. the creation of new forms (as proposed by Mair 2006). In IndE, it is the combination of a modal auxiliary and the progressive that accounts for most complex forms, while in PhiE, the distribution of the complex forms is more even. In JamE, SinE and HKE, on the other hand, the proportion of fragmentary forms of the progressive is considerably higher than in any other variety. Overall, the progressive in World Englishes is found in present tense progressives co-occurring with *am, are, is* and their contracted forms.

The lexical variation attested in World Englishes is rather subtle, based on the present results. In most varieties, the most frequently used progressive main verbs are the same, i.e. *going, doing* and *saying*. However, the most frequent progressive main verbs in HKE differ to a great extent from the main verbs used in other varieties, most likely due to the data included in ICE-HK. The distribution of the semantic domains that the progressive occurs in follows the restrictions laid out by the semantic properties of the progressive: verbs of activity and communication account for close to 75% of all progressives in the present data. As regards IC and OC varieties, the only major difference was found in the amount of mental verbs: OC varieties use mental verbs clearly more often than IC varieties. The investigation of progressive and non-progressive main verbs provided more evidence to support the finding: the verbs preferring the progressive are activity or communication verbs, such as KID and JOKE, while verbs dispreferring the progressive are most often mental verbs, such as LOVE and KNOW. The lexical variation of the progressive is closely related to the functions of the progressive, as

at times, it is the main verb that triggers the function: this is the case with stative progressives, in particular.

The main findings related to the functions of the progressive in World Englishes are the following: first, subjective progressives are the largest group of progressives referring to situations other than the prototypically progressive ones, and there is a significant difference between IC and OC varieties (IC varieties having a higher frequency of subjective progressives). Second, stative progressives are equally frequent in both IC and OC varieties, but it is the extended use of stative verbs that sets the OC varieties apart from the IC varieties. IndE and HKE, in particular, use stative progressives to refer to situations that are not perceived as temporary and are thus incompatible with the progressive in StE. Third, the progressive is used to refer to situations in the future clearly more often in the OC varieties, although on the level of individual varieties, we find that most frequent use of futurate progressives is found in SinE, IndE and IrE. Fourth, extended use of habitual progressives is found in IndE and HKE, similarly to extended use of stative progressives. It may thus be concluded that extended use of the progressive is not generalizable to OC varieties overall, but is rather restricted to individual varieties, i.e. IndE and HKE. Various previous studies have already indicated that IndE is using the progressive in an innovative way, but the result concerning HKE is somewhat surprising. More information could be gained by investigating the aspectual system of the substrate languages in more detail, or by applying theories of second language acquisition to HKE.

In order to summarize on the use of the progressive in individual varieties, Table 55 below presents the findings of this study. Additionally, the second column provides information on the use of the progressive in World Englishes in general. As Table 55 portrays, the use of the progressive in individual varieties differs to a great extent, starting from the frequency of the construction and ending with the distribution of the functions. IrE in particular seems to function differently in many aspects, for instance, the distribution of tense forms is very different from any other variety. Similarly, the functions of the progressive in IrE are distributed more evenly: BrE and AmE are much more heavily tilted towards the prototypical uses. The fact that IrE differs to such an extent from the other two varieties categorized here as IC varieties suggests that another theoretical model might be more appropriate to depict the situation between these three varieties. What exactly this model should be is to be discussed in further studies.

Table 55. Summary of the uses of the progressive in World Englishes.

	WEs	BrE	IrE	AmE
Frequency (RF)	6,345	842	1,073	771
Present tense	71%	71%	51%	59%
Past tense	29%	29%	49%	41%
SimpProg	87%	91%	90%	89%
ModalProg	5%	2%	5%	3%
PerfProg	3%	4%	2%	5%
PassProg	1%	1%	<1%	<1%
Fragmentary	3%	1%	2%	3%
Main clause PF	67%	60%	73%	70%
Contracted PF	43%	55%	42%	48%
Temporal ADVs	18%	15%	19%	14%
Main verb	*going* (9%)	*doing* (10%)	*going* (15%)	*doing* (10%)
50% cut-off	13 verbs	11	9	14
Prototypical	39%	52%	29%	42%
Subjective	18%	10%	26%	23%
Stative	14%	13%	19%	16%
Futurate	12%	9%	17%	5%
Habitual	8%	9%	4%	11%
Other + Indet.	9%	8%	5%	3%

JamE	IndE	PhiE	SinE	HKE
859	890	802	733	375
79%	86%	71%	73%	87%
21%	14%	29%	27%	13%
82%	84%	88%	87%	85%
4%	12%	5%	4%	3%
4%	2%	4%	3%	3%
2%	<1%	2%	<1%	<1%
8%	2%	<1%	6%	9%
58%	74%	66%	66%	66%
53%	20%	49%	37%	40%
18%	17%	21%	21%	19%
doing (12%)	*doing* (8%)	*going* (9%)	*going* (11%)	*studying* (9%)
11	13	16	11	11
42%	36%	37%	38%	34%
23%	8%	24%	13%	14%
11%	18%	14%	14%	11%
6%	16%	11%	19%	11%
10%	6%	8%	8%	9%
8%	16%	5%	9%	21%

As regards IndE, the eye catches some peculiarities compared to other varieties - other OC varieties as well as all other varieties: the proportion of modal auxiliaries combined with the progressive is significantly higher, while the proportions of contracted forms, or subjective progressives is much lower than in any other variety. What is most notable about HKE, on the other hand, is the fact that regardless of the very low frequency, the progressive is used in a manner very close to the World Englishes average. Of course, there are exceptions, such as the extended use of stative and habitual progressives, found also in IndE. Overall, it seems that the varieties differing most from the other varieties are IrE and IndE, as well as HKE to some extent.[158] Similarly, Schilk & Hammel (2014: 164) found that, in their investigation into stative progressives in the dialogue section of some ICE-corpora, IndE and HKE were the ones to differ from all other varieties.

To conclude, this study set out to compare the use of the progressive in altogether eight varieties of English spoken the world over: the goal of the study has now been reached, and the result is a description of the progressive in World Englishes. Although the amount of previous research on the progressive is considerable, the present study has added the worldwide dimension to the ongoing discussion. Only in the past few years is research being focused on the use of the progressive in varieties other than BrE or AmE: the study is thus located at the core of current interest.

9.2 Methodological considerations

Undoubtedly, some of the methodological choices made in the course of the study are open to criticism. For instance, the focus on spoken data only, rather than a comparison of progressives occurring in speech and writing, led to problems as regards the frequency of features that are more prominent in written language, such as the progressive passive. Furthermore, most previous studies focus on written data, or both written and spoken, which means that comparison to results obtained in these previous studies was, at times, problematic. However, the choice to concentrate on spoken data is also easily defended: the progressive is most frequently used in spoken language (as shown in section 4.2). Further, including a

[158] Noting down the variety or varieties that differ from the others separately for each row in Table 55, it becomes evident that IrE is the variety noted down most often (9 times), with IndE and HKE following (6 and 5 times, respectively).

comparison of results from spoken and written data would have made the study even more complicated than it is now.

Another possible problem related to the data is its size. At times, it was evident that the 100,000 words per variety is not enough to produce generalizable results. For instance, the more fine-grained analysis of types of subordinate clauses, or matters related to the progressive passive faced difficulties due to the amount of data. The problem is, in part, related to the focus on spoken data: features which are more frequent in written language require a substantial amount of spoken data in order to return a sufficient number of hits. Further problems due to the restricted amount of data was encountered in relation to statistical testing. The expected frequency of any cell should be more than 5: this requirement was not always fulfilled without restructuring of the data.

The double perspective taken up in the present study, i.e. the study of the progressive in isolation and as a member of the progressive vs. non-progressive paradigm, proved to be a fruitful decision. Owing to the fact that making comparisons to the non-progressive requires a considerable amount of time, not all linguistic features were included in the comparison. However, on most occasions, investigation into the progressive vs. non-progressive paradigm provided valuable additional information not only on the progressive, but also related to spoken English and World Englishes in general. For instance, the discussion on the use of modal auxiliaries with the progressive gained insight and depth from the comparison of the non-progressive paradigm. Although at the beginning of the study, the progressive vs. non-progressive paradigm was added to the study of the frequency of the progressive, the comparison proved to be more interesting in the case of morphosyntactic variation of the progressive in World Englishes.

9.3 Suggestions for further research

Throughout the study, we have encountered matters that call for further research, as they, albeit being interesting, fall outside the scope of the study. The investigation of the progressive vs. non-progressive paradigm unveiled a need for further research regarding grammatical features other than the progressive, such as passive constructions or expressions of modality in World Englishes. Regarding the progressive in World Englishes, the limitations of the present data could be avoided by adding more spoken data to the data set, or by adding the written components of the ICE. The comparison of spoken and written genres has already

produced interesting results in studies such as Biber et al. (1999), Leech et al. (2009), and regarding World Englishes, Collins (2008). In particular, the study of the complex forms of the progressive could shed light on the channels of spread in World Englishes: the present study indicates that, at least in spoken language, the progressive has spread in World Englishes via the simple forms. The study of written data could verify this result and suggest other explanations. In addition to the complex progressives, another feature worth investigating in detail is the futurate progressive. The present data indicates a difference in the use of futurate progressive in IC and OC varieties: a larger set of data could provide answers as to why OC varieties, and SinE in particular, use futurate progressives so frequently.

A considerably more large-scale endeavor is related to the investigation of the diachronic changes of the progressive in World Englishes. Diachronic data is available on BrE and AmE[159], but the fact that diachronic corpora representing other World Englishes are practically non-existent makes the task difficult, if not impossible to accomplish. Researchers developing diachronic corpora on World Englishes are facing a number of problems, not the least of all being lack of material: documents written in English in the colonies are difficult to locate. Even if such historical texts are acquired, further questions arise regarding, for instance, whose texts should be included and whose rejected.[160] Despite these obvious difficulties in compiling diachronic corpora representing World Englishes, the need for such corpora is not diminished. Investigation into the historical development of the progressive in World Englishes would provide valuable information on the progressive, and perhaps more importantly, on World Englishes.

[159] For instance: The Helsinki Corpus of English Texts on BrE, and The Corpus of Historical American English (COHA) on AmE.

[160] Sebastian Hoffmann, personal communication, April 4th, 2014.

References

Corpora and corpus tools

The *International Corpus of English: Great Britain*. 1998. The Survey of English Usage. University College London.

The *International Corpus of English: Hong Kong*. 2006. Department of English, The University of Hong Kong and The Chinese University of Hong Kong, PRC.

The *International Corpus of English: India*. 2002. Department of English, Shivaji University, Kolhapur and Freie Universität, Berlin.

The *International Corpus of English: Ireland*. 2008. School of English, The Queen's University Belfast and Centre for Language and Communication Studies, Trinity College Dublin.

The *International Corpus of English: Jamaica*. 2009. Englisches Seminar, Albert-Ludwigs-Universität Freiburg and Department of Language, Linguistics, & Philosophy, University of the West Indies.

The *International Corpus of English: the Philippines*. 2004. Department of English and Applied Linguistics, De La Salle University, Manila.

The *International Corpus of English: Singapore*. 2002. Department of English Language & Literature, The National University of Singapore.

The *Santa Barbara Corpus of Spoken American English*. 2002-2005. Santa Barbara Center for the Study of Discourse, University of California, Santa Barbara.

Scott, Mike. 2012. WordSmith Tools version 6. Liverpool: Lexical Analysis Software.

Bibliography

Aarts, Bas, Joanne Close, Geoffrey Leech & Sean Wallis (eds.). 2013. *The Verb Phrase in English. Investigating Recent Language Change with Corpora*. Cambridge: Cambridge University Press.

Ansaldo, Umberto. 2004. The evolution of Singapore English: Finding the matrix. In Lisa Lim (ed.), *Singapore English: A Grammatical Description*, 127-149. Amsterdam: John Benjamins Publishing Company.

Arnaud, René. 1998. The development of the progressive in 19th-century English: A quantitative survey. *Language Variation and Change* 10, 123-152.

Balasubramanian, Chandrika. 2009. *Register Variation in Indian English*. Amsterdam: John Benjamins Publishing Company.

Bao, Zhiming & Hong Huaqing. 2006. Diglossia and register variation in Singapore English. *World Englishes* 25(1), 105-114.

Barry, Michael. 1982. The English language in Ireland. In Richard W. Bailey & Manfred Görlach (eds.), *English as a World Language,* 84-133. Cambridge: Cambridge University Press.

Bean, Frank & Gillian Stevens. 2003. *America's Newcomers and the Dynamics of Diversity*. New York: Russell Sage Foundation.
Bhatt, Rakesh M. 2008. *Indian English: Syntax*. In Mesthrie (ed.), 546-562.
Biber, Douglas, Stig Johansson, Geoffrey Leech, Susan Conrad & Edward Finegan. 1999. *Longman Grammar of Spoken and Written English*. Harlow, Essex: Pearson Education Ltd.
Binnick, Robert I. 1991. *Time and the Verb. A Guide to Tense and Aspect*. Oxford: Oxford University Press.
Bolt, Philip & Kingsley Bolton. 1996. The International Corpus of English in Hong Kong. In Greenbaum (ed.), 197-214.
Bolton, Kingsley (ed.). 2002. *Hong Kong English. Autonomy and Creativity*. Hong Kong: Hong Kong University Press.
Bolton, Kingsley. 2003. *Chinese Englishes: A Sociolinguistic History*. Cambridge: Cambridge University Press.
Bolton, Kingsley & David Graddol. 2012. English in China today. *English Today* 28(3): 3-9.
Brinton, Laurel J. 1988. *The Development of English Aspectual Systems: Aspectualizers and Post-verbal Particles*. Cambridge: Cambridge University Press.
Bybee, Joan, Revere Perkins & William Pagliuca. 1994. *The Evolution of Grammar: Tense, Aspect, and Modality in the Languages of the World*. Chicago/London: The University of Chicago Press.
Carter, Ronald & Michael McCarthy. 2006. *Cambridge Grammar of English. A Comprehensive Guide. Spoken and Written English Grammar and Usage*. Cambridge: Cambridge University Press.
Cheng, Chin-Chuan. 1992. Chinese varieties of English. In Braj Kachru (ed.), *The Other Tongue. English across Cultures,* 2nd edn, 162-177. Urbana, Chicago, London: University of Illinois Press.
Cheshire, Jenny (ed.). 1991. *English around the World. Sociolinguistic Perspectives*. Cambridge: Cambridge University Press.
Collins, Peter. 2008. The progressive aspect in World Englishes: A corpus-based study. *Australian Journal of Linguistics* 28(2), 225-249.
Collins, Peter. 2009. The progressive. In Pam Peters, Peter Collins & Adam Smith (eds.), *Comparative Studies in Australian and New Zealand English. Grammar and beyond*, 115-123. Philadelphia: John Benjamins Publishing Company.
Collins, Peter & Xinyue Yao. 2012. Modals and quasi-modals in New Englishes. In Hundt & Gut (eds.), 35-53.
Comrie, Bernard. 1976. A*spect: An Introduction to the Study of Verbal Aspect and Related Problems*. Cambridge: Cambridge University Press.
Couper-Kuhlen, Elizabeth. 1995. On the foregrounded progressive in American conversational narrative: A new development? In Wolfgang Riehle and Hugo Keiper (eds.), *Anglistentag 1994 Graz*. Tübingen: Niemeyer, 229-245.
Crystal, David. 1966. Specification and English tenses. *Journal of Linguistics* 2, 1-34.
Crystal, David. 2003. *English as a Global Language*, 2nd edn. Cambridge: Cambridge University Press.
Davydova, Julia. 2012. Englishes in the Outer and Expanding Circles: A comparative study. *World Englishes* 31(3), 366-385.
Declerck, Renaat. 1991. *Tense in English: its structure and use in discourse*. London: Routledge.
Depraetere, Ilse & Susan Reed. 2000. The present perfect progressive: constraints on its use with numerical object NPs. *English Language and Linguistics* 4(1), 97-114.

Deterding, David. 2007. *Singapore English*. Edinburgh: Edinburgh University Press.
Deuber, Dagmar. 2009. 'The English we speaking': Morphological and syntactic variation in educated Jamaican speech. *Journal of Pidgin and Creole Languages* 24(1), 1-52.
Deuber, Dagmar. 2014. *English in the Caribbean: Variation, Style and Standards in Jamaica and Trinidad*. Cambridge: Cambridge University Press.
Devonish, Hubert & Otelemate G. Harry. 2008. Jamaican Creole and Jamaican English: phonology. In Schneider (ed.), 256-289.
Dixon, R. M. W. 2005. *A Semantic Approach to English Grammar,* 2nd edn. Oxford: Oxford University Press.
Elsness, Johan. 1994. On the progression of the progressive in early Modern English. *ICAME Journal* 18, 5-25.
Filppula, Markku. 1999. *The Grammar of Irish English: Language in Hibernian style*. London: Routledge.
Filppula, Markku. 2003. More on the English progressive and the Celtic connection. In Hildegard Tristram (ed.), *The Celtic Englishes III*, 150-168. Heidelberg: Universitätsverlag C. Winter.
Filppula, Markku. 2008. Irish English: morpohology and syntax. In Kortmann & Upton (eds.), 328-359.
Filppula, Markku, Juhani Klemola & Heli Paulasto (eds.). 2009. *Vernacular Universals and Language Contacts: Evidence from Varieties of English and Beyond*. New York/London: Routledge.
Filppula, Markku & Juhani Klemola. 2012. English in contact: Celtic and Celtic Englishes. In Alex Bergs & Laurel Brinton (eds.), *Historical Linguistics of English* (HSK 34.2), 1687-1703. Berlin/Boston: de Gruyter.
Gachelin, Jean-Marc. 1997. The progressive and habitual aspects in non-standard Englishes. In Edgar W. Schneider (ed.), *Englishes around the World: Studies in honour of Manfred Görlach. Volume 1: General studies, British Isles, North America*, 33-46. Amsterdam: John Benjamins Publishing Company.
Gonzalez, Andrew. 2004. The social dimensions of Philippine English. *World Englishes* 23(1), 7-16.
Görlach, Manfred. 1991. *Englishes: Studies in Varieties of English 1984-1988*. Amsterdam: John Benjamins Publishing Company.
Gramley, Stephan & Kurt-Michael Pätzold. 2004. *A Survey of Modern English,* 2nd edn. London: Routledge.
Greenbaum, Sidney (ed.). 1996. *Comparing English Worldwide. The International Corpus of English*. Oxford: Clarendon Press.
Gupta, Anthea F. 1994. *The Step-tongue: Children's English in Singapore*. Clevedon, UK: Multilingual Matters.
Harris, John. 1991a. Conservatism versus substratal transfer in Irish English. In Peter Trudgill & J.K. Chambers (eds.), *Dialects of English: Studies in Grammatical Variation*, 191-211. London: Longman.
Harris, John. 1991b. Ireland. In Cheshire (ed.), 37-50.
Hickey, Raymond. 2007. *Irish English. History and present-day forms*. Cambridge: Cambridge University Press.
Hickey, Raymond. 2008. Irish English: phonology. In Kortmann & Upton (eds.), 71-104.
Hilbert, Michaela & Manfred Krug. 2012. Progressives in Maltese English. A comparison with spoken and written text types of British and American English. In Hundt & Gut (eds.), 101-136.

Hoffmann, Thomas & Lucia Siebers (eds.). 2009. *World Englishes - Problems, Properties and Prospects. Selected papers from the 13th IAWE conference.* Amsterdam: John Benjamins Publishing Company.

Huddleston, Rodney & Geoffrey K. Pullum. 2002. *The Cambridge Grammar of the English Language*. Cambridge: Cambridge University Press.

Hundt, Marianne. 2004a. The passival and the progressive passive. A case study of layering in the English aspect and voice systems. In Hans Lindquist (ed.), *Corpus Approaches to Grammaticalization in English*, 79-120. Philadelphia: John Benjamins Publishing Company.

Hundt, Marianne. 2004b. Animacy, agentivity, and the spread of the progressive in Modern English. *English Language and Linguistics* 8(1), 47-69.

Hundt, Marianne. 2009a. Global feature - local norms? A case study on the progressive passive. In Hoffmann & Siebers (eds.), 287-308.

Hundt, Marianne. 2009b. Colonial lag, colonial simplification, or simply language change? In Günther Rohdenburg & Julia Schlüter (eds.), *One Language, Two Grammars? Differences between British and American English*, 13-37. Cambridge: Cambridge University Press.

Hundt, Marianne. 2013. The diversification of English: old, new and emerging epicentres. In Daniel Schreier & Marianne Hundt (eds.), *English as a Contact Language*, 182-203. Cambridge: Cambridge University Press.

Hundt, Marianne & Katrin Vogel. 2011. Overuse of the progressive in ESL and learner Englishes - fact or fiction? In Joybrato Mukherjee & Marianne Hundt (eds.), *Exploring Second-Language Varieties of English and Learner Englishes: Bridging a Paradigm Gap*, 145-165. Amsterdam: John Benjamins Publishing Company.

Hundt, Marianne & Ulrike Gut (eds.). 2012. *Mapping Unity and Diversity World-Wide. Corpus-based Studies of New Englishes.* Amsterdam: John Benjamins Publishing Company.

Hung, Tony T.N. 2012. Hong Kong English. In Ee-Ling Low & Azirah Hashim (eds.), *English in Southeast Asia: Features, policy and language in use*, 113-133. Amsterdam: John Benjamins.

Jenkins, Jennifer. 2003. *World Englishes. A resource book for students*. London: Routledge.

Kachru, Braj B. 1985. Standards, codification and sociolinguistic realism: the English language in the outer circle. In Randolph Quirk & H.G. Widdowson (eds.), *English in the World. Teaching and Learning the Language and Literature*, 11-30. Cambridge: Cambridge University Press.

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)*. Cló Ollscoil na Banríona.

Kilgarriff, Adam. 2005. Language is Never, Ever, Ever, Random. *Corpus Linguistics and Linguistic Theory* 1, 263-275.

Klein, Wolfgang. 1994. *Time in language*. London: Routledge.

Kortmann, Bernd & Clive Upton (eds.). 2008. *Varieties of English* vol.1: *The British Isles*. Berlin/New York: Mouton de Gruyter.

Kortmann, Bernd & Kerstin Lunkenheimer (eds.). 2013. *The electronic World Atlas of Varieties of English* [eWAVE]. Leipzig: Max Planck Institute for Evolutionary Anthropology. Available online at: http://www.ewave-atlas.org/.

Kranich, Svenja. 2010. *Progressive in Modern English: A Corpus-Based Study of Grammaticalization and Related Changes*. Amsterdam: Rodopi.

Leech, Geoffrey N. 1987. *Meaning and the English Verb*, 2nd edn. London/New York: Longman.
Leech, Geoffrey & Jan Svartvik. 1994. *A Communicative Grammar of English,* 2nd edn. Harlow, Essex: Longman Group Ltd.
Leech, Geoffrey, Marianne Hundt, Christian Mair & Nicholas Smith. 2009. *Change in Contemporary English*. Cambridge: Cambridge University Press.
Leith, Dick. 2007. English - colonial to postcolonial. In David Graddol, Dick Leith, Joan Swann, Martin Rhys & Julia Gillen (eds.), *Changing English*, 2nd edn, 117-152. London: Routledge.
Levin, Magnus. 2013. The progressive verb in modern American English. In Aarts et al. (eds.), 187-216.
Lim, Lisa. 2009. Not just an "Outer Circle", "Asian" English: Singapore English and the significance of ecology. In Hoffmann & Siebers (eds.), 180-205.
Ljung, Magnus. 1980. *Reflections on the English Progressive*. Gothenburg Studies in English 46. Acta Universitatis Gothoburgensis.
Mair, Christian. 2006. *Twentieth-Century English: History, Variation, and Standardization*. Cambridge: Cambridge University Press.
Mair, Christian. 2009. Corpus linguistics meets sociolinguistics. Studying educated spoken usage in Jamaica on the basis of the International Corpus of English. In Hoffmann & Siebers (eds.), 39-60.
Mair, Christian. 2011. Corpora and the New Englishes. Using the 'Corpus of Cyber-Jamaican' to explore research perspective for the future. In Meunier, Fanny, Gaëtanelle Gilquin & Magali Paquot (eds.), *A Taste for Corpora. In Honour of Sylviane Granger*, 209-236. Amsterdam: John Benjamins Publishing Company.
Mair, Christian. 2013. The World System of Englishes: Accounting for the transnational importance of mobile and mediated vernaculars. *English World-Wide* 34(3), 253-278.
Mair, Christian & Marianne Hundt. 1995. Why is the Progressive Becoming More Frequent in English. A Corpus-Based Investigation of Language Change in Progress. *Zeitschrift für Anglistik und Amerikanistik* 43(2), 111-122.
Marckwardt, Albert H. 1958. *American English*. New York: Oxford University Press.
McKay, Sandra Lee. 2010. English as an International Language. In Nancy H. Hornberger & Sandra Lee McKay (eds.), *Sociolinguistics and Language Education*, 89-115. Clevedon: Multilingual Matters.
Meriläinen, Lea, Heli Paulasto & Paula Rautionaho. Forthcoming. Extended use of the progressive in Inner Circle, Outer Circle and Extended Circle varieties (working title). To appear in Markku Filppula, Juhani Klemola & Anna Mauranen (eds.), *Changing English: Global and Local Perspectives [working title]*.
Mesthrie, Rajend (ed.). 2008. *Varieties of English* vol. 4: *Africa, South and Southeast Asia*. Berlin, New York: Mouton de Gruyter.
Mesthrie, Rajend & Rakesh M. Bhatt (eds.). 2008. *World Englishes: The Study of New Linguistic Varieties*. Cambridge: Cambridge University Press.
Mufwene, Salikoko. 2001. *The Ecology of Language Evolution*. Cambridge: Cambridge University Press.
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.
Nelson, Gerard, Sean Wallis & Bas Aarts. 2002. *Exploring Natural Language: Working with the British Component of the International Corpus of English*. Amsterdam: John Benjamins Publishing Company.

Nesselhauf, Nadja. 2007. The spread of the progressive and its 'future' use. *English Language and Linguistics* 11(1), 191-207.
Nesselhauf, Nadja. 2009. Co-selection phenomena across New Englishes. Parallels (and differences) to foreign learner varieties. *English World-Wide* 30(1), 1-26.
Nesselhauf, Nadja & Ute Römer. 2007. Lexical-grammatical patterns in spoken English. The case of the progressive with future time reference. *International Journal of Corpus Linguistics* 12(3), 297-333.
Patrick, Peter L. 2008. Jamaican Creole: morphology and syntax. In Schneider (ed.), 609-644.
Paulasto, Heli. 2014. Extended uses of the progressive in L1 and L2 Englishes. *English World-wide* 35(3), 247-276.
Peters, Pam, Peter Collins & Adam Smith (eds.). 2002. *New Frontiers of Corpus Research: Papers from the 21st International Conference on English Language Research in Computerized Corpora, Sydney 2002*. Amsterdam: Rodopi.
Pfaff, Meike, Alexander Bergs & Thomas Hoffmann. 2013. *I was just reading this article* – on the expression of recentness and the English past progressive. In Aarts et al. (eds.), 217-238.
Platt, John. 1991. Social and linguistic constraints on variation in the use of two grammatical variables in Singapore English. In Cheshire (ed.), 376-387.
Platt, John, Heidi Weber & Ho Mian Lian. 1984. *The New Englishes*. London: Routledge & Kegan Paul.
Quirk, Randolph, Sidney Greenbaum, Geoffrey Leech & Jan Svartvik. 1985. *A Comprehensive Grammar of the English Language*. London/New York: Longman.
Rickford, John. 1999. Variation in the Jamaican Creole copula and its relation to the genesis of AAVE: New data and analysis. In John Rickford & Suzanne Romaine (eds.), *Creole genesis, attitudes and discourse: Studies celebrating Charlene J. Sato*, 143-156. Amsterdam: Benjamins.
Rogers, Chandrika K. 2002. Syntactic features of Indian English. An examination of written Indian English. In Randi Reppen, Susan M. Fitzmaurice & Douglas Biber (eds.), *Using Corpora to Explore Linguistic Variation*, 187-202. Philadelphia: John Benjamins Publishing Company.
Römer, Ute. 2005. *Progressives, Patterns, Pedagogy: A Corpus-Driven Approach to English Progressive Forms, Functions, Contexts and Didactics*. Philadelphia: John Benjamins Publishing Company.
Sahgal, Anju. 1991. Patterns of language use in a bilingual setting in India. In Cheshire (ed.), 299-306.
Sailaja, Pingali. 2009. *Indian English*. Edinburgh: Edinburgh University Press.
Scheffer, Johannes. 1975. *The Progressive in English*. Amsterdam: North-Holland Publishing Company.
Schilk, Marco & Marc Hammel. 2014. The progressive in South Asian and Southeast Asian varieties of English – mapping areal homogeneity and heterogeneity. In Lieven Vandelanotte, Kristin Davidse & Caroline Gentens (eds.), *Recent Advances in Corpus Linguistics: Developing and Exploiting Corpora*, 147-171. Amsterdam: Rodopi.
Schlüter, Norbert. 2002. Temporal specification of the present perfect: a corpus-based study. In Peters et al. (eds.), 307-316.
Schmied, Josef. 1990. Corpus linguistics and non-native varieties of English. *World Englishes* 9(3), 255-268.

Schneider, Edgar W. 2003. The dynamics of New Englishes: From identity construction to dialect birth. *Language* 79(2), 233-281.
Schneider, Edgar W. 2007. *Postcolonial English. Varieties around the world.* Cambridge: Cambridge University Press.
Schneider, Edgar W. 2014. New reflections on the evolutionary dynamics of world Englishes. *World Englishes* 33(1), 9-32.
Schneider, Edgar W. (ed.). 2008. *Varieties of English*, vol. 2: *The Americas and the Caribbean*. Berlin, New York: Mouton de Gruyter.
Sebba, Mark. 1997. *Contact Languages: Pidgins and Creoles.* Basingstoke: Macmillan.
Sedlatschek, Andreas. 2009. *Contemporary Indian English.* Amsterdam: John Benjamins Publishing Company.
Sharma, Devyani. 2001. The pluperfect in native and non-native English: A comparative corpus study. *Language Variation and Change* 13, 343-373.
Sharma, Devyani. 2009. Typological diversity in New Englishes. *English World-Wide* 30(2), 170-195.
Smith, Carlota S. 1991. *The Parameter of Aspect*. The Netherlands: Kluwer Academic Publishers.
Smith, Nicholas. 2002. Ever moving on?: The progressive in recent British English. In Peters et al. (eds.), 317-330.
Smith, Nicholas Ian. 2005. *A Corpus-based Investigation of Recent Change in the Use of the Progressive in British English*. PhD dissertation, Lancaster University.
Smith, Nicholas & Paul Rayson. 2007. Recent change and variation in the British English use of the progressive passive. *ICAME Journal* 31, 129-160.
Smitterberg, Erik. 2005. *The Progressive in 19th-century English: A Process of Integration*. Amsterdam: Rodopi.
Smitterberg, Erik, Sabine Reich & Angela Hahn. 2000. The present progressive in political and academic language in the 19th and 20th centuries: A corpus-based investigation. *ICAME Journal* 24, 99-118.
Strang, Barbara. 1970. *A History of English*. London/New York: Methuen.
Strang, Barbara. 1982. Some aspects of the history of the *be+ing* construction. In John Anderson (ed.), *Language Form and Linguistic Variation. Papers dedicated to Angus McIntosh*, 427-474. Amsterdam: John Benjamins Publishing Company.
Suoniemi, Paula. 2012. Variation in the progressive in World Englishes: Some preliminary findings. In Sebastian Hoffmann, Paul Rayson & Geoffrey Leech (eds.), *English Corpus Linguistics: Looking back, Moving forward: Papers from the 30th International Conference on English Language Research on Computerized Corpora (ICAME 30). Lancaster, UK, 27-31 May 2009*, 205-215. Amsterdam/New York: Rodopi.
Svartvik, Jan & Geoffrey Leech. 2006. *English: One Tongue, Many Voices.* Basingstoke: Palgrave Macmillan.
Tay, Mary W. J. 1991. Southeast Asia and Hong Kong. In Cheshire (ed.), 319-332.
Tayao, Ma. Lourdes G. 2008. Philippine English: phonology. In Mesthrie (ed.), 292-306.
Thompson, Roger M. 2003. *Filipino English and Taglish. Language Switching from Multiple Perspectives.* Amsterdam: Benjamins.
Tottie, Gunnel. 2002. *An Introduction to American English*. Oxford: Blackwell Publishing.
Trudgill, Peter. 2004. *New-Dialect Formation: The Inevitability of Colonial Englishes*. Edinburgh: Edinburgh University Press.
Trudgill, Peter & Jean Hannah. 2002. *International English: A Guide to the varieties of Standard English*, 4th edn. London: Edward Arnold.

Tsui, Amy B. M. & David Bunton. 2002. The discourse and attitudes of English language teachers in Hong Kong. In Bolton, 57-77.
Van Rooy, Bertus. 2006. The extension of the progressive aspect in Black South African English. *World Englishes* 25(1), 37-64.
Van Rooy, Bertus. 2014. Progressive aspect and stative verbs in Outer Circle varieties. *World Englishes* 33(2), 157-172.
Vendler, Zeno. 1957. Verbs and Times. *The Philosophical Review* 66(2), 143-160.
Webster, Jonathan J. 2009. Language in Hong Kong. In Hoffmann & Siebers (eds.), 143-153.
Wee, Lionel. 2008. Singapore English: morphology and syntax. In Mesthrie (ed.), 593-609.
Williams, Christopher. 2002. *Non-Progressive and Progressive Aspect in English*. Fasano: Schena Editore.
Wolfram, Walt & Natalie Schilling-Estes. 2006. *American English: Dialects and Variation,* 2nd edn. Cambridge/Oxford: Basil Blackwell.
Yip, Virginia & Stephen Matthews. 2000. *Basic Cantonese: Grammar and Workbook*. London: Routledge.
Ziegeler, Debra. 2006. *Interfaces with English Aspect: Diachronic and Empirical Studies*. Amsterdam: John Benjamins Publishing Company.

Appendix A

Search results for *are* + V*ing* in all categories of ICE-India. Note that the number of progressives given has been rounded, and that the results should be considered tentative only. The distribution of *are* + V*ing* indicates that the progressive is clearly more frequent in the spoken categories, and in private dialogues, in particular.

Categories	Number of files	Progressives (RF)	Progressives (per file)
Private dialogues	100	520	5.2
Public dialogues	80	370	4.6
Unscripted monologues	70	200	2.9
Scripted monologues	50	130	2.6
Written categories	200	220	1.1

Appendix B

Design of the ICE corpora: http://ice-corpora.net/ice/design.htm.

SPOKEN (300)	**Dialogues** (180)	**Private** (100)	Face-to-face conversations (90) Phonecalls (10)
		Public (80)	Classroom Lessons (20) Broadcast Discussions (20) Broadcast Interviews (10) Parliamentary Debates (10) Legal cross-examinations (10) Business Transactions (10)
	Monologues (120)	**Unscripted** (70)	Spontaneous commentaries (20) Unscripted Speeches (30) Demonstrations (10) Legal Presentations (10)
		Scripted (50)	Broadcast News (20) Broadcast Talks (20) Non-broadcast Talks (10)
WRITTEN (200)	**Non-printed** (50)	**Student Writing** (20)	Student Essays (10) Exam Scripts (10)
		Letters (30)	Social Letters (15) Business Letters (15)
	Printed (150)	**Academic writing** (40)	Humanities (10) Social Sciences (10) Natural Sciences (10) Technology (10)
		Popular writing (40)	Humanities (10) Social Sciences (10) Natural Sciences (10) Technology (10)
		Reportage (20)	Press news reports (20)
		Instructional writing (20)	Administrative Writing (10) Skills/hobbies (10)
		Persuasive writing (10)	Press editorials (10)
		Creative writing (20)	Novels & short stories (20)

Appendix C

Files included in the present study.

Corpus	Files
ICE-GB	S1A-001, S1A-003, S1A-005, S1A-006, S1A-007, S1A-009, S1A-011, S1A-013, S1A-014, S1A-015, S1A-017, S1A-019, S1A-021, S1A-022, S1A-023, S1A-025, S1A-027, S1A-029, S1A-031, S1A-033, S1A-035, S1A-036, S1A-037, S1A-039, S1A-041, S1A-043, S1A-045, S1A-047, S1A-048, S1A-049, S1A-051, S1A-053, S1A-055, S1A-056, S1A-057, S1A-059, S1A-061, S1A-063, S1A-064, S1A-065, S1A-067, S1A-069, S1A-070, S1A-071, S1A-073, S1A-075, S1A-076, S1A-077, S1A-079, S1A-081, S1A-083, S1A-084, S1A-085, S1A-087, S1A-089
SBCSAE	SBC001, SBC002, SBC004, SBC005, SBC007, SBC011, SBC015, SBC017, SBC019, SBC031, SBC032, SBC036, SBC043, SBC044, SBC047, SBC049, SBC050, SBC051, SBC056, SBC058, SBC060
ICE-IRL	S1A-001, S1A-002, S1A-004, S1A-006, S1A-007, S1A-008, S1A-010, S1A-013, S1A-014, S1A-016, S1A-019, S1A-020, S1A-022, S1A-025, S1A-026, S1A-028, S1A-031, S1A-032, S1A-034, S1A-037, S1A-038, S1A-040, S1A-043, S1A-044, S1A-046, S1A-047, S1A-049, S1A-050, S1A-052, S1A-055, S1A-056, S1A-058, S1A-061, S1A-062, S1A-064, S1A-067, S1A-068, S1A-070, S1A-073, S1A-074, S1A-076, S1A-079, S1A-080, S1A-082, S1A-083, S1A-085, S1A-086, S1A-088
ICE-JA	S1A-001, S1A-002, S1A-004, S1A-006, S1A-007, S1A-008, S1A-010, S1A-013, S1A-014, S1A-016, S1A-019, S1A-020, S1A-022, S1A-025, S1A-026, S1A-028, S1A-031, S1A-032, S1A-034, S1A-037, S1A-038, S1A-040, S1A-044, S1A-046, S1A-047, S1A-048, S1A-049, S1A-052, S1A-053, S1A-055, S1A-058, S1A-059, S1A-061, S1A-064, S1A-065, S1A-067, S1A-070, S1A-071, S1A-073, S1A-076, S1A-077, S1A-079, S1A-082, S1A-083, S1A-085, S1A-088, S1A-089
ICE-IND	S1A-001, S1A-002, S1A-004, S1A-006, S1A-007, S1A-008, S1A-010, S1A-013,

	S1A-014, S1A-016, S1A-019, S1A-020, S1A-022, S1A-025, S1A-026, S1A-028, S1A-031, S1A-032, S1A-034, S1A-037, S1A-038, S1A-040, S1A-043, S1A-044, S1A-046, S1A-047, S1A-049, S1A-050, S1A-052, S1A-055, S1A-056, S1A-058, S1A-061, S1A-062, S1A-064, S1A-067, S1A-068, S1A-070, S1A-073, S1A-074, S1A-076, S1A-079, S1A-080, S1A-082, S1A-085, S1A-086, S1A-088
ICE-PHI	S1A-001, S1A-003, S1A-005, S1A-006, S1A-007, S1A-009, S1A-011, S1A-013, S1A-015, S1A-017, S1A-019, S1A-021, S1A-023, S1A-025, S1A-027, S1A-029, S1A-031, S1A-033, S1A-035, S1A-037, S1A-039, S1A-041, S1A-043, S1A-045, S1A-047, S1A-049, S1A-051, S1A-053, S1A-055, S1A-057, S1A-059, S1A-061, S1A-063, S1A-065, S1A-067, S1A-069, S1A-071, S1A-073, S1A-075, S1A-077, S1A-079, S1A-081, S1A-083, S1A-085, S1A-087, S1A-089
ICE-SIN	S1A-001, S1A-003, S1A-005, S1A-006, S1A-007, S1A-009, S1A-011, S1A-013, S1A-015, S1A-017, S1A-019, S1A-021, S1A-023, S1A-025, S1A-027, S1A-029, S1A-031, S1A-033, S1A-035, S1A-037, S1A-039, S1A-041, S1A-043, S1A-045, S1A-047, S1A-048, S1A-049, S1A-051, S1A-053, S1A-055, S1A-057, S1A-059, S1A-061, S1A-063, S1A-065, S1A-067, S1A-069, S1A-071, S1A-073, S1A-075, S1A-077, S1A-079, S1A-081, S1A-083, S1A-085, S1A-087, S1A-089
ICE-HK	S1A-001, S1A-002, S1A-003, S1A-004, S1A-005, S1A-013, S1A-014, S1A-015, S1A-016, S1A-017, S1A-018, S1A-025, S1A-026, S1A-027, S1A-028, S1A-029, S1A-030, S1A-037, S1A-038, S1A-039, S1A-040, S1A-041, S1A-042, S1A-050, S1A-051, S1A-052, S1A-053, S1A-061, S1A-062, S1A-063, S1A-065, S1A-073, S1A-074, S1A-075, S1A-076, S1A-077, S1A-086, S1A-087, S1A-088, S1A-089, S1A-090

Appendix D

Search strings for progressives, tagged.

Search string	Retrieves	Example with the verb GIVE.
*PR	progressive, present	*is giving*
*PRM	progressive, present, modal aux	*may be giving*
*PRFM	progressive, present, future, modal aux	*will be giving*
*PRF	progressive, present, future	*is giving* (futurity evident from context)
*PA	progressive, past	*was giving*
*PAM	progressive, past, modal aux	*might be giving*
*PAFM	progressive, past, future, modal aux	
*PAF	progressive, past, future	*was giving* (futurity evident from context)
*FRAG	progressive, BE omitted	*Ø giving*

Search strings for non-progressives, tagged.

Search string	Retrieves:	Example with the verb GIVE
*VR	non-progressive, present	*gives*
*VRM	non-progressive, present, modal aux	*may give*
*VRFM	non-progressive, present, future, modal aux	*will give*
*VRF	non-progressive, present, future	*gives* (futurity evident from context)
*VA	non-progressive, past	*gave*
*VAM	non-progressive, past, modal aux	*might give*
*VAFM	non-progressive, past, future, modal aux	
*VAF	non-progressive, past, future	**was going to**

Appendix E

Appendix E presents the chi-square tests for statistical significance, sorted out according to the table or figure on which the calculations are based. The chi-square tests are performed on the overall distribution, on the distribution of the feature in IC and OC varieties, and on two varieties at a time: the varieties are ordered according to the frequency of the feature in focus, and the statistical testing is performed on two varieties located next to each other. For instance, the frequency of feature X is the lowest in variety A and second lowest in variety B, and so forth; statistical testing is then performed on varieties A+B, B+C, C+D, etc. Note: * indicates a probability of less than .05, ** a probability of less than .01, and *** a probability of less than .001.

Table 18 and Figure 6. Frequency of the progressive in World Englishes, M-coefficient.

	Progressives	Other words in the corpus
BrE	842	99,568
IrE	1,073	98,955
AmE	771	99,237
JamE	859	99,652
IndE	890	100,176
PhiE	802	100,074
SinE	733	100,079
HKE	375	100,137

Overall distribution: X-squared = 349.9006, df = 7, p-value < 2.2e-16 ***
Distribution in IC vs. OC: X-squared = 67.6132, df = 1, p-value < 2.2e-16 ***
HKE vs. SinE: X-squared = 115.2415, df = 1, p-value < 2.2e-16 ***
SinE vs. AmE: X-squared = 1.2982, df = 1, p-value = 0.2545
AmE vs. PhiE: X-squared = 0.3754, df = 1, p-value = 0.5401
PhiE vs. BrE: X-squared = 1.1769, df = 1, p-value = 0.278
BrE vs. JamE: X-squared = 0.1545, df = 1, p-value = 0.6942
JamE vs. IndE: X-squared = 0.3954, df = 1, p-value = 0.5295
IndE vs. IrE: X-squared = 19.1894, df = 1, p-value = 1.184e-05 ***

Table 19 and Figure 7. Frequency of the progressive in World Englishes, V-coefficient.

	Progressive VPs	Non-progressive VPs
BrE	842	13,728
IrE	1,073	12,480
AmE	771	13,577
JamE	859	11,811
IndE	890	11,348
PhiE	802	12,237
SinE	733	12,422
HKE	375	10,041

Overall distribution: X-squared = 253.5852, df = 7, p-value < 2.2e-16 ***
Distribution in IC vs. OC: X-squared = 6.2151, df = 1, p-value = 0.01267 *(*)
HKE vs. AmE: X-squared = 43.0001, df = 1, p-value = 5.474e-11 ***
AmE vs. SinE: X-squared = 0.5229, df = 1, p-value = 0.4696
SinE vs. BrE: X-squared = 0.5527, df = 1, p-value = 0.4572
BrE vs. PhiE: X-squared = 1.6984, df = 1, p-value = 0.1925
PhiE vs. JamE: X-squared = 4.2071, df = 1, p-value = 0.04026 *
JamE vs. IndE: X-squared = 2.314, df = 1, p-value = 0.1282
IndE vs. IrE: X-squared = 3.8005, df = 1, p-value = 0.05124 (*)

Figure 8. Distribution of present and past tense simple progressive VPs in World Englishes.

	Present tense simple progressive VPs	Past tense simple progressive VPs
BrE	544	222
IrE	495	475
AmE	405	279
JamE	552	149
IndE	638	107
PhiE	492	212
SinE	464	178
HKE	275	43

Overall distribution: X-squared = 360.319, df = 7, p-value < 2.2e-16 ***
Distribution in IC vs. OC: X-squared = 217.3522, df = 1, p-value < 2.2e-16 ***
IrE vs. AmE: X-squared = 10.8197, df = 1, p-value = 0.001004 ***
AmE vs. PhiE: X-squared = 17.2961, df = 1, p-value = 3.198e-05 ***
PhiE vs. BrE: X-squared = 0.2259, df = 1, p-value = 0.6346
BrE vs. SinE: X-squared = 0.6457, df = 1, p-value = 0.4216
SinE vs. JamE: X-squared = 6.1244, df = 1, p-value = 0.01333 **
JamE vs. IndE: X-squared = 11.7778, df = 1, p-value = 0.0005994 ***
IndE vs. HKE: X-squared = 0.1299, df = 1, p-value = 0.7186

Figure 9. Distribution of the progressive with and without modal auxiliaries in World Englishes.

	Progressive VPs with modal auxiliaries	Progressive VPs without modal auxiliaries
BrE	20	822
IrE	56	1,017
AmE	23	748
JamE	35	824
IndE	106	784
PhiE	43	759
SinE	32	701
HKE	10	365

Overall distribution: X-squared = 112.2851, df = 7, p-value < 2.2e-16 ***
Distribution in IC vs. OC: X-squared = 19.7737, df = 1, p-value = 8.717e-06 ***
BrE vs. HKE: X-squared = 0.0916, df = 1, p-value = 0.7621
HKE vs. AmE: X-squared = 0.0904, df = 1, p-value = 0.7637
AmE vs. JamE: X-squared = 1.4102, df = 1, p-value = 0.235
JamE vs. SinE: X-squared = 0.0831, df = 1, p-value = 0.7731
SinE vs. IrE: X-squared = 0.6842, df = 1, p-value = 0.4081
IrE vs. PhiE: X-squared = 0.0187, df = 1, p-value = 0.8914
PhiE vs. IndE: X-squared = 22.5267, df = 1, p-value = 2.072e-06 ***

Figure 10. Distribution of perfect and non-perfect progressives in World Englishes.

	Perfect progressive VPs	Non-perfect progressive VPs
BrE	34	808
IrE	24	1,049
AmE	37	734
JamE	34	825
IndE	18	872
PhiE	31	771
SinE	18	715
HKE	12	363

Overall distribution: X-squared = 18.9492, df = 7, p-value = 0.008348 **
Distribution in IC vs. OC: X-squared = 0.983, df = 1, p-value = 0.3215
IndE vs. IrE: X-squared = 0.1067, df = 1, p-value = 0.744
IrE vs. SinE: X-squared = 0.0919, df = 1, p-value = 0.7618
SinE vs. HKE: X-squared = 0.5218, df = 1, p-value = 0.4701
HKE vs. PhiE: X-squared = 0.3214, df = 1, p-value = 0.5708
PhiE vs. JamE: X-squared = 0.0095, df = 1, p-value = 0.9224
JamE vs. BrE: X-squared = 0.0071, df = 1, p-value = 0.933
BrE vs. AmE: X-squared = 0.5538, df = 1, p-value = 0.4568

Figure 11. Perfect progressives occurring with temporal adverbials.

	Perfect progressive VPs with temporal adverbial	Perfect progressive VPs without temporal adverbial
BrE	13	21
IrE	12	12
AmE	14	23
JamE	15	19
IndE	6	12
PhiE	17	14
SinE	3	15
HKE	5	7

Overall distribution: X-squared = 8.5072, df = 6, p-value = 0.2032 (HKE excluded, expected value is less than 5)
Distribution in IC vs. OC: X-squared = 0.0025, df = 1, p-value = 0.9598

Figure 12. Distribution of passive and active progressive VPs in World Englishes.

	Passive progressive VPs	Active progressive VPs
BrE	11	831
IrE	5	1,068
AmE	4	767
JamE	18	841
IndE	5	885
PhiE	19	783
SinE	6	727
HKE	2	373

Overall distribution: X-squared = 28.216, df = 6, p-value = 8.556e-05 *** (HKE excluded, expected value is less than 5)
Distribution in IC vs. OC: X-squared = 5.4906, df = 1, p-value = 0.01912 *
IrE vs. AmE: expected value less than 5
AmE vs. HKE: expected value less than 5
HKE vs. IndE: expected value less than 5
IndE vs. SinE: expected value less than 5
SinE vs. BrE: X-squared = 0.8735, df = 1, p-value = 0.35
BrE vs. JamE: X-squared = 1.5797, df = 1, p-value = 0.2088
JamE vs. PhiE: X-squared = 0.1426, df = 1, p-value = 0.7057

Table 35. Distribution of simple, complex and fragmentary progressives in World Englishes.

	Simple progressive VPs	Complex progressive VPs	Fragmentary progressive VPs
BrE	766	65	11
IrE	970	85	18
AmE	684	64	23
JamE	701	87	71
IndE	745	129	16
PhiE	704	93	5
SinE	636	56	41
HKE	318	24	33

Overall distribution: X-squared = 194.4246, df = 14, p-value < 2.2e-16 ***
Distribution in IC vs. OC: X-squared = 46.994, df = 2, p-value = 6.243e-11 ***
JamE vs. IndE: X-squared = 43.7399, df = 2, p-value = 3.177e-10 ***
IndE vs. HKE: X-squared = 47.726, df = 2, p-value = 4.329e-11 ***
HKE vs. SinE: X-squared = 4.4589, df = 2, p-value = 0.1076
SinE vs. PhiE: X-squared = 37.7873, df = 2, p-value = 6.231e-09 ***
PhiE vs. AmE: X-squared = 16.6118, df = 2, p-value = 0.0002471 ***
AmE vs. IrE: X-squared = 3.6613, df = 2, p-value = 0.1603
IrE vs. BrE: X-squared = 0.4708, df = 2, p-value = 0.7903

Table 36. Distribution of simple, complex and fragmentary progressive verb phrases in Inner and Outer Circle Englishes.

	Inner Circle	Outer Circle
SimpProg	2,420	3,104
PerfProg	95	113
ModalProg	99	226
PassProg	20	50
FragProg	52	166

Overall distribution: X-squared = 60.568, df = 4, p-value = 2.204e-12 ***

Table 37. Distribution of the progressive in main and subordinate clauses in World Englishes (subordinate as one class, unclear cases excluded).

	Progressive VPs in main clauses	Progressive VPs in subordinate clauses
BrE	504	333
IrE	778	280
AmE	540	218
JamE	499	356
IndE	662	215
PhiE	532	269
SinE	482	247
HKE	247	123

Overall distribution: X-squared = 101.8708, df = 7, p-value < 2.2e-16 ***
Distribution in IC vs. OC: X-squared = 2.7741, df = 1, p-value = 0.0958
JamE vs. BrE: X-squared = 0.6013, df = 1, p-value = 0.4381
BrE vs. SinE: X-squared = 5.822, df = 1, p-value = 0.01583 *
SinE vs. PhiE: X-squared = 0.0153, df = 1, p-value = 0.9017
PhiE vs. HKE: X-squared = 0.0131, df = 1, p-value = 0.9088
HKE vs. AmE: X-squared = 2.3695, df = 1, p-value = 0.1237
AmE vs. IrE: X-squared = 1.1685, df = 1, p-value = 0.2797
IrE vs. IndE: X-squared = 0.9574, df = 1, p-value = 0.3278

Table 39. Contracted and non-contracted progressive VPs in World Englishes (fragmentary forms excluded).

	Contracted progressive VPs	Non-contracted progressive VPs
BrE	464	367
IrE	454	601
AmE	368	380
JamE	458	330
IndE	175	699
PhiE	394	403
SinE	269	423
HKE	147	195

Overall distribution: X-squared = 339.0546, df = 7, p-value < 2.2e-16 ***
Distribution in IC vs. OC: X-squared = 34.3031, df = 1, p-value = 4.716e-09 ***
IndE vs. SinE: X-squared = 67.5544, df = 1, p-value < 2.2e-16 ***
SinE vs. HKE: X-squared = 1.6076, df = 1, p-value = 0.2048
HKE vs. IrE: X-squared = 3e-04, df = 1, p-value = 0.9869
IrE vs. AmE: X-squared = 6.7055, df = 1, p-value = 0.009612 **
AmE vs. PhiE: X-squared = 0.0087, df = 1, p-value = 0.9256
PhiE vs. JamE: X-squared = 12.0269, df = 1, p-value = 0.0005244 ***
JamE vs. BrE: X-squared = 0.8617, df = 1, p-value = 0.3533

Figure 13. Distribution of progressive VPs with and without temporal adverbials in World Englishes.

	Progressive VPs with temporal adverbials	Progressive VPs without temporal adverbials
BrE	123	719
IrE	199	874
AmE	109	661
JamE	157	702
IndE	154	736
PhiE	171	631
SinE	153	580
HKE	72	303

Overall distribution: X-squared = 25.3386, df = 7, p-value = 0.0006607 ***
Distribution in IC vs. OC: X-squared = 11.2497, df = 1, p-value = 0.0007964 ***
AmE vs. BrE: X-squared = 0.0668, df = 1, p-value = 0.7961
BrE vs. IndE: X-squared = 2.3395, df = 1, p-value = 0.1261
IndE vs. JamE: X-squared = 0.2835, df = 1, p-value = 0.5944
JamE vs. IrE: X-squared = 0.023, df = 1, p-value = 0.8795
IrE vs. HKE: X-squared = 0.0781, df = 1, p-value = 0.7799
HKE vs. SinE: X-squared = 0.4291, df = 1, p-value = 0.5124
SinE vs. PhiE: X-squared = 0.0463, df = 1, p-value = 0.8297

Figure 15. Different types of temporal adverbials modifying progressives in World Englishes.

	Time-reference	Frequency	Duration
BrE	88	11	24
IrE	145	15	39
AmE	71	6	32
JamE	110	16	31
IndE	115	12	27
PhiE	103	17	51
SinE	108	10	35
HKE	48	8	16

Overall distribution: X-squared = 17.1171, df = 14, p-value = 0.25
Distribution in IC vs. OC: X-squared = 0.9166, df = 2, p-value = 0.6324
PhiE vs. AmE: X-squared = 1.8578, df = 2, p-value = 0.395
AmE vs. HKE: X-squared = 2.61, df = 2, p-value = 0.2712
HKE vs. JamE: X-squared = 0.27, df = 2, p-value = 0.8737
JamE vs. SinE: X-squared = 1.594, df = 2, p-value = 0.4507
SinE vs. BrE: X-squared = 0.8889, df = 2, p-value = 0.6412
BrE vs. IrE: X-squared = 0.2045, df = 2, p-value = 0.9028
IrE vs. IndE: X-squared = 0.2441, df = 2, p-value = 0.8851

Table 44. Distribution of the semantic domains in Inner and Outer Circle varieties.

	Inner Circle	Outer Circle
Activity	374	448
Communication	127	179
Mental	59	136
Existence	57	66
Occurrence	29	46
Aspectual	8	13
Causative	9	12
Indeterminate	10	17

Overall distribution: X-squared = 16.8009, df = 7, p-value = 0.01873 *

Figure 19. Functions of the progressive in Inner and Outer Circle Englishes.

	Inner Circle	Outer Circle
Prototypical	307	470
Subjective	148	202
Stative	121	169
Futurate	78	157
Habitual	59	104
Other + Indeterminate	37	148

Overall distribution: X-squared = 33.1214, df = 5, p-value = 3.56e-06 ***

Figure 20. Prototypical progressives in World Englishes.

	Prototypical progressives	Other progressives
BrE	129	121
IrE	73	177
AmE	105	145
JamE	105	145
IndE	91	159
PhiE	93	157
SinE	95	155
HKE	86	164

Overall distribution: X-squared = 32.073, df = 7, p-value = 3.937e-05 ***
Distribution in IC vs. OC: X-squared = 2.1924, df = 1, p-value = 0.1387
IrE vs. HKE: X-squared = 1.5585, df = 1, p-value = 0.2119
HKE vs. IndE: X-squared = 0.2186, df = 1, p-value = 0.6401
IndE vs. PhiE: X-squared = 0.0344, df = 1, p-value = 0.8529
PhiE vs. SinE: X-squared = 0.0341, df = 1, p-value = 0.8535
SinE vs. AmE: X-squared = 0.8333, df = 1, p-value = 0.3613
AmE vs. JamE: X-squared = 0, df = 1, p-value = 1
JamE vs. BrE: X-squared = 4.627, df = 1, p-value = 0.03147 *

Figure 21. Subjective uses of the progressive in World Englishes.

	Subjective progressives	Other progressives
BrE	26	224
IrE	65	185
AmE	57	193
JamE	58	192
IndE	19	231
PhiE	59	191
SinE	32	218
HKE	34	216

Overall distribution: X-squared = 61.6035, df = 7, p-value = 7.217e-11 ***
Distribution in IC vs. OC: X-squared = 4.1457, df = 1, p-value = 0.04174 *
IndE vs. BrE: X-squared = 1.1966, df = 1, p-value = 0.274
BrE vs. SinE: X-squared = 0.7021, df = 1, p-value = 0.4021
SinE vs. HKE: X-squared = 0.0698, df = 1, p-value = 0.7916
HKE vs. AmE: X-squared = 7.1066, df = 1, p-value = 0.00768 **
AmE vs. JamE: X-squared = 0.0113, df = 1, p-value = 0.9154
JamE vs. PhiE: X-squared = 0.0112, df = 1, p-value = 0.9159
PhiE vs. IrE: X-squared = 0.3861, df = 1, p-value = 0.5344

Figure 23. Stative uses of the progressive in World Englishes.

	Stative progressives	Other progressives
BrE	33	217
IrE	48	202
AmE	40	210
JamE	27	223
IndE	44	206
PhiE	36	214
SinE	34	216
HKE	28	222

Overall distribution: X-squared = 12.3089, df = 7, p-value = 0.09085
Distribution in IC vs. OC: X-squared = 2.5822, df = 1, p-value = 0.1081
JamE vs. HKE: X-squared = 0.0204, df = 1, p-value = 0.8863
HKE vs. BrE: X-squared = 0.4668, df = 1, p-value = 0.4945
BrE vs. SinE: X-squared = 0.0172, df = 1, p-value = 0.8956
SinE vs. PhiE: X-squared = 0.0664, df = 1, p-value = 0.7966
PhiE vs. AmE: X-squared = 0.2483, df = 1, p-value = 0.6183
AmE vs. IndE: X-squared = 0.2289, df = 1, p-value = 0.6323
IndE vs. IrE: X-squared = 0.2131, df = 1, p-value = 0.6443

Figure 24. Extended and standard stative uses of the progressive in World Englishes.

	Extended stative progressives	Standard stative progressives
BrE	1	32
IrE	4	44
AmE	3	37
JamE	4	23
IndE	16	28
PhiE	2	34
SinE	3	31
HKE	9	19

Overall distribution: Expected values less than 5 for Extended statives in BrE, JamE, SinE and HKE.
Distribution in IC vs. OC: X-squared = 10.3866, df = 1, p-value = 0.001269 **

Figure 25. Futurate uses of the progressive in World Englishes.

	Futurate progressives	Other progressives
BrE	23	227
IrE	43	207
AmE	12	238
JamE	14	236
IndE	41	209
PhiE	28	222
SinE	47	203
HKE	27	223

Overall distribution: X-squared = 46.9797, df = 7, p-value = 5.633e-08 ***
Distribution in IC vs. OC: X-squared = 2.1091, df = 1, p-value = 0.1464
AmE vs. JamE: X-squared = 0.1623, df = 1, p-value = 0.6871
JamE vs. BrE: X-squared = 2.3641, df = 1, p-value = 0.1242
BrE vs. HKE: X-squared = 0.3556, df = 1, p-value = 0.551
HKE vs. PhiE: X-squared = 0.0204, df = 1, p-value = 0.8863
PhiE vs. IndE: X-squared = 2.8414, df = 1, p-value = 0.09186
IndE vs. IrE: X-squared = 0.0572, df = 1, p-value = 0.8109
IrE vs. SinE: X-squared = 0.2168, df = 1, p-value = 0.6415

Figure 26. Subcategories of futurate progressives.

	Will/shall + be Ving	Futurity evident from context
BrE	1	22
IrE	12	31
AmE	3	9
JamE	2	12
IndE	12	29
PhiE	7	21
SinE	11	36
HKE	8	19

Overall distribution: Expected values less than 5 for *will/shall* in AmE and JamE.
Distribution in IC vs. OC: X-squared = 0.7077, df = 1, p-value = 0.4002

Figure 28. Habitual uses of the progressive in World Englishes.

	Habitual progressives	Other progressives
BrE	22	228
IrE	9	241
AmE	28	222
JamE	25	225
IndE	16	234
PhiE	21	229
SinE	19	231
HKE	23	227

Overall distribution: X-squared = 12.8176, df = 7, p-value = 0.07668
Distribution in IC vs. OC: X-squared = 0.1287, df = 1, p-value = 0.7198
IrE vs. IndE: X-squared = 2.0632, df = 1, p-value = 0.1509
IndE vs. SinE: X-squared = 0.2765, df = 1, p-value = 0.599
SinE vs. PhiE: X-squared = 0.1087, df = 1, p-value = 0.7416
PhiE vs. BrE: X-squared = 0.0254, df = 1, p-value = 0.8733
BrE vs. HKE: X-squared = 0.0244, df = 1, p-value = 0.8758
HKE vs. JamE: X-squared = 0.0922, df = 1, p-value = 0.7614
JamE vs. AmE: X-squared = 0.1899, df = 1, p-value = 0.663

Figure 29. Extended and standard habitual progressives.

	Extended habitual progressives	Standard habitual progressives
BrE	1	21
IrE	0	9
AmE	1	27
JamE	5	20
IndE	7	9
PhiE	2	19
SinE	1	18
HKE	5	18

Overall distribution: Expected values less than 5 for extended habitual throughout.
Distribution in IC vs. OC: X-squared = 8.0908, df = 1, p-value = 0.004449 **

Table 53. Functions of the progressive in World Englishes.

	Prototyp.	Subj.	Stat.	Futur.	Habit.	Other+Indet.
BrE	129	26	33	23	22	17
IrE	73	65	48	43	9	12
AmE	105	57	40	12	28	8
JamE	105	58	27	14	25	21
IndE	91	19	44	41	16	39
PhiE	93	59	36	28	21	13
SinE	95	32	34	47	19	23
HKE	86	34	28	27	23	52

Overall distribution: X-squared = 202.6408, df = 35, p-value < 2.2e-16 ***
Distribution in IC vs. OC: X-squared = 33.1214, df = 5, p-value = 3.56e-06 ***

Table 54. Frequency of aspectual, extended and subjective progressives in World Englishes.

	Aspectual	Extended	Subjective	Other+Indet
BrE	205	2	26	17
IrE	169	4	65	12
AmE	181	4	57	8
JamE	162	9	58	21
IndE	169	23	19	39
PhiE	174	4	59	13
SinE	191	4	32	23
HKE	150	14	34	52

Overall distribution: X-squared = 176.2781, df = 21, p-value < 2.2e-16 ***
Distribution in IC vs. OC: X-squared = 43.333, df = 3, p-value = 2.091e-09 ***