2018

Revisiting the ENL-ESL-EFL continuum: A multifactorial approach to grammatical aspect in spoken Englishes

Rautionaho, Paula

Walter de Gruyter GmbH

Revisiting the ENL-ESL-EFL continuum: A multifactorial approach to grammatical aspect in spoken Englishes

Paula Rautionaho[1], Sandra C. Deshors[2] and Lea Meriläinen[1]
University of Eastern Finland[1], Michigan State University[2]

Abstract

This study focuses on the progressive vs. non-progressive alternation to revisit the debate on the ENL-ESL-EFL continuum (i.e. whether native (ENL) and non-native (ESL/EFL) Englishes are dichotomous types of English or form a gradient continuum). While progressive marking is traditionally studied independently of its unmarked counterpart, we examine (i) how the grammatical contexts of both constructions systematically affect speakers' constructional choices in ENL (American, British), ESL (Indian, Nigerian and Singaporean) and EFL (Finnish, French and Polish learner Englishes) and (ii) what light speakers' varying constructional choices bring to the continuum debate. Methodologically, we use a clustering technique to group together individual varieties of English (i.e. to identify similarities and differences between those varieties) based on linguistic contextual features such as AKTIONSART, ANIMACY, SEMANTIC DOMAIN (of aspect-bearing lexical verb), TENSE, MODALITY and VOICE to assess the validity of the ENL-ESL-EFL classification for our data. Then, we conduct a logistic regression analysis (based on lemmas observed in both progressive and non-progressive constructions) to explore how grammatical contexts influence speakers' constructional choices differently across English types. While, overall, our cluster analysis supports the ENL-ESL-EFL classification as a useful theoretical framework to explore cross-variety variation, the regression shows that, when we start digging into the specific linguistic contexts of (non-)progressive constructions, this classification does not systematically transpire in the data in a uniform manner. Ultimately, by including more than one statistical technique into their exploration of the continuum, scholars could avoid potential methodological biases.

1   Introduction

Research on varieties of English has until recently traditionally been anchored in Kachru's (1985) Three Circles classification model of Englishes, which equates native English (ENL), English as a Second Language (ESL) and English as a Foreign Language (EFL) with the status of English in a particular country – ENL or Inner Circle countries, where English is spoken as the dominant native language (e.g. UK, US), ESL or Outer Circle countries, where English has spread as a result of colonialism and serves as an additional official language (e.g. India, Singapore) and EFL or Expanding Circle countries, where English is learnt as an important foreign language in formal instructional settings (e.g., France, Japan). Despite the fact that the Three Circles model has started to receive some criticism (mainly because it is based on colonial heritage and does not reflect the globalization process that Englishes have undergone over the past few years; see Mair 2013), it has nonetheless, in the recent years, provided a valuable theoretical framework to explore whether ENL, ESL and EFL form a continuum rather than strict categories (see Hundt and Mukherjee 2011). Even though, traditionally, World Englishes scholars have assessed the characteristics of ESL varieties mainly against native yardsticks like British English (henceforth BrE), as a result of the "Second Language Varieties of English and Learner Englishes" 2008 workshop on the occasion of the *First Conference of the International Society for the Study of English*, a number of scholars started to integrate EFL into their analyses of World Englishes and to explore the three types of Englishes in a "unified" fashion (Deshors 2014; also see Mukherjee and Hundt eds. 2011, Gilquin 2015, Gries and Deshors 2015 for studies that adopt such a unified analytical framework).

By bringing ENL, EFL and ESL together, researchers have been able to identify parallels in the uses of English by foreign and second language speakers (in the broad sense of the term; see Mukherjee and Hundt eds. 2011, Edwards and Laporte 2015, Gilquin 2015, Gries and Deshors 2015, Meriläinen and Paulasto 2017). Therefore, studies such as Deshors (2014), Edwards (2014, 2016) and Gilquin and Granger (2011) have called into question the traditionally assumed divide between EFL and ESL by showing that it is not necessarily clear-cut. For instance, Deshors (2014: 298) concludes that "within the EFL–ESL continuum, individual world and learner variants are intermingled rather than grouped together according to 'type' and positioned distinctively closer or further away from the native variant". In this context, a tendency common to a large number of corpus-based studies on the continuum is that they start by analyzing linguistic phenomena in individual English varieties that belong to differ-

ent types of Englishes (e.g. Hong Kong, Singapore, Indian English for ESL, German, French, Spanish learner English for EFL and British, American, Australian English for ENL, etc.). From there, observed differences across these varieties serve as a springboard to draw generalized conclusions on the structure of ENL, ESL and EFL at large. In this context, what the present study is set up to do is to distinguish itself from these studies in the following way: first, by contrasting individual Englishes from all three types of varieties (using a clustering technique) to verify that the varieties we investigate yield patterns of language use compatible with Kachru's classification. Then, in a second step, we contribute to the continuum discussion by integrating the three types of English as a predictor of potential linguistic variation directly into a logistic regression model so as to assess to what extent linguistic patterns can be reliably predicted to belong to a specific type of English. While neither of our methodological steps is new in itself (both techniques were applied independently in Szmrecsanyi and Kortmann (2011) and Edwards and Laporte (2015) for the cluster analysis and Gries and Deshors (2015) for the regression analysis), the innovative aspect of our approach lies in that we use both techniques to explore the same linguistic phenomenon and we use the regression approach to dig deeper into the results of the cluster analysis.[1] This is an important point because the cluster and the regression analyses differ significantly in how they model and ultimately portray the usage patterns of a particular linguistic phenomenon. For instance, in the cluster analysis, potential interactions between the use of a particular linguistic feature and the elements of its linguistic context of use are not taken into account, whereas they are accounted for in the regression analysis. Ultimately, it is reasonable to believe that accounting for the context of use of a particular linguistic feature and its potential influence on how that feature is being used by different populations of English speakers may lead to different conclusions on the validity of the continuum.

For our purposes, we specifically focus on progressive marking in two ENLs (British (BrE) and American (AmE) Englishes), three ESLs (Indian (IndE), Singaporean (SgE) and Nigerian (NigE) Englishes) and three EFL varieties of English (by Finnish-, French- and Polish-speaking learners). Although this linguistic feature has long been explored as a way to capture cross-varietal variation (e.g. van Rooy 2006, 2014; Hundt and Vogel 2011; Rautionaho 2014; van Rooy and Piotrowska 2015; Laitinen and Levin 2016; Edwards 2016 and Deshors 2017), it is only recently that it has been studied as part of the progressive vs. non-progressive alternation. Specifically, in the first multifactorial analysis of progressive vs. non-progressive constructions, Rautionaho and Deshors's (forthc.) study on the linguistic contexts of the two constructions to determine

whether and to what extent those contexts influence the constructional choices of speakers of different ESL/EFL varieties. Although the study adds illuminating findings to existing monofactorial accounts of the progressive construction, a main downside of the study is its limitation to written data. This is important in that the progressive has been shown to be more frequent in speech than in writing (e.g. Leech et al. 2009, Salles Bernal 2015). Furthermore, according to Salles Bernal (2015), the combination of the progressive with certain linguistic features (e.g. the modal auxiliary *will*) is not only more frequent in speech than in writing but specifically so in AsianEnglish varieties while the opposite is true of BrE (2015: 98). Against this body of research, the question arises as to how speakers (rather than writers) of different types of Englishes differ in their choices of progressive vs. non-progressive construction and how such potential differences can inform the overall debate on the ENL-ESL-EFL continuum.

In what follows, we first set the stage for the present study by discussing previous research on the ENL-ESL-EFL continuum and on the progressive (Section 2), after which we present the data and the statistical methods in Section 3. The results of the study are presented in Section 4, and they are discussed in detail in Section 5.

2 Setting the stage

2.1 Why is Kachru's classification of Englishes a matter of debate for the continuum?

As mentioned above, the Three Circles model, and the ENL-ESL-EFL distinction, is anchored in the historical context of English varieties. As the Kachruvian paradigm draws attention to and seeks acceptance of post-colonial varieties (i.e. ESLs) and their local standards and norms, it is primarily the different sociolinguistic contexts of EFLs and ESLs that set them apart. It is assumed that those different contexts (wider society for ESL vs. classrooms for EFL) are likely to give rise to different orientation to norms: in this framework, ESL is viewed as norm-*developing* whereas EFL is viewed as norm-*dependent* (Kachru 1985). Put differently, societies where English is spoken as an ESL provide ample opportunities for speakers to use the language, which results in the spread and conventionalization of non-standard forms within the speech community (van Rooy 2010, 2011, Deshors et al. 2016). In EFL contexts, however, non-standard forms are generally equated with 'deficiency' (Deshors 2017; see Deshors et al. 2016 for an in-depth discussion on errors vs. innovations in the speech of EFL and ESL users). A number of recent corpus-based studies have begun to empirically examine the validity of the labels ESL and EFL in the context of the continuum

on the grounds that the two non-native types of Englishes arise from similar processes of second language acquisition (e.g. Williams 1987, Biewer 2011, Meriläinen and Paulasto 2017) and also because the global spread of English has led many EFL speakers to take their use of English beyond the classroom walls, thus blurring the sociolinguistic divide between EFL and ESL (Mair 2013). In this context, studies such as Bongartz and Buschfeld (2011) on the post-colonial variety of Cyprus English and Gilquin and Granger (2011) on the use of *into* (which has been found to trigger innovative uses in non-native varieties), support the existence of the continuum (rather than categorical distinction) between ESL and EFL. More recent, and methodologically more sophisticated studies such as Edwards and Laporte (2015) on *into* across AmE, BrE, SgE, IndE and Dutch English and Deshors (2014) on the dative alternation across BrE, IndE, SgE, Hong Kong English (HKE) and German and French learner Englishes also support the existence and the theoretical relevance of the continuum. Interestingly, however, while more methodologically sophisticated studies such as Deshors (2014), Gries and Deshors (2015) and Szmrecsanyi and Kortmann (2011) demonstrate the usefulness of state-of-the-art statistical techniques such as (mixed-effects) logistic regression modeling (in the case of Deshors 2014, and Gries and Deshors 2015,) and cluster analysis (in the case of Szmrecsanyi and Kortmann 2011,) to contribute to the continuum discussion, those three studies draw somewhat conflicting conclusions as to whether or not the distinction between EFL and ESL should be viewed as continuous or dichotomous in nature. While, on the one hand, Deshors (2014) supports the continuum (i.e. a non-categorical approach to ESL and EFL), Gries and Deshors (2015) as well as Szmrecsanyi and Kortmann (2011), on the other hand, lean more towards a dichotomous distinction. For instance, based on a large-scale comparison of synthetic vs. analytic coding strategies in grammar, Szmrecsanyi and Kortmann (2011) claim the need for drawing a distinction between EFL and ESL (see also Laitinen (forthc.) in the context of English as a Lingua Franca).

The present work builds on the above studies by continuing to apply sophisticated statistical techniques but also by integrating Kachru's classification directly into a data modeling process, something that, as far as we know, neither of the above-mentioned studies include in their methodological design. A main benefit of this approach lies in that it allows us to investigate the same linguistic phenomenon, progressive marking, through the lens of both techniques, cluster analysis and logistic regression, within the scope of a single analysis and this, with the view to assess to what extent both statistical approaches, can ultimately lead to differing perspectives on whether or not EFL and ESL should be located on a continuum.

2.2 Why is progressive marking helpful to talk about the continuum?

A number of studies have already shown that World Englishes vary in their uses of the progressive. Studies such as Hundt and Vogel (2011), Rautionaho (2014), van Rooy (2006, 2014), van Rooy and Piotrowska (2015), and Laitinen and Levin (2016) have all explored the uses of the progressive form (in particular in terms of its frequencies and semantic functions) within individual English varieties and integrated their findings into the general question of the continuum. As a linguistic phenomenon, progressive marking is interesting for various reasons. First, over the past few centuries, its frequency has increased in spoken language and speech-like registers in particular (e.g. Smitterberg 2005; Leech et al. 2009: 119–141). Second, previous studies show that the progressive is indeed used in different functions and linguistic contexts, although the differences are commonly attested in individual varieties rather than types of Englishes (see e.g. Collins 2008; Hundt and Vogel 2011; Rautionaho 2014). Furthermore, semantically, ESL varieties are often characterized by the extended use of stative and habitual progressives in situations that are not temporary in nature (see, e.g. van Rooy 2006, 2014; van Rooy and Piotrowska 2015). Importantly, though, there is considerable variation between ESL varieties; extended stative or habitual progressives are infrequent in SgE and Philippines English and relatively common in IndE and HKE (Rautionaho 2014: 199, 207). Turning to contexts of use, or the morpho-syntax of the progressive, Collins (2008), for instance, shows that SgE and IndE use progressives with modal auxiliaries considerably more frequently than ENL and other ESL varieties, while Rautionaho (2014: 104) found that HKE and IndE speakers combine the progressive with the present tense clearly more often than ENL and other ESL speakers.

Recently, the analytical focus in research on the progressive has started to shift to the progressive vs. non-progressive alternation (rather than focusing on the analysis of the progressive as an isolated, stand-alone construction) with the aim of pinning down the contextual factors that influence speakers' choice of one linguistic variant or the other. Importantly, research on the alternation is primarily interested in the linguistic contexts of use of the two constructions, and it does not address the differences or similarities of the functions of the two constructions. See Aarts et al. (2013) for discussions on how to measure the increase in the progressive – per million words or in comparison with non-progressives – and to what extent there are identifiable 'true alternants'. Rautionaho and Deshors (forthc.) present a logistic regression analysis of the two constructions in BrE, AmE, IndE, SgE and Dutch English varieties, and find that, more often than not, more than one factor simultaneously affects the choice of the progressive over the non-progressive. More specifically, semantic domains emerge

as contextual features that influence writers' constructional choices regardless of their English variety and their written genre. Furthermore, Existence verbs, Aspect-Causative verbs and Mental verbs most significantly influence writers' choices, and as those domains are shown to transcend varieties and genres, they emerge as core determining factors in writers' constructional choices.[2] In the context of the continuum, the study supports the recent trend to explore the fuzzy boundaries of ESL and EFL in progressive marking (Hundt and Vogel 2011; Meriläinen et al. 2017) in that both AmE and Dutch English were found to differ from BrE (with regard to past tense use in AmE and modal uses in Dutch English), while the investigated ESL varieties did not differ from BrE. Against this background, Rautionaho and Deshors (forthc.) addressed an existing gap in the literature. In this context, the present study builds on the above body of research by exploring:

(i)   to what extent individual English varieties on the ENL-ESL-EFL continuum yield different alternation patterns of progressive vs. non-progressive constructions?

(ii) which linguistic contextual factors contribute to those different alternation patterns?

(iii) to what extent spoken data (as opposed to written data) help us further our understanding of progressive marking in World Englishes?

3   Data and statistical approach

3.1   English varieties under investigation and corpus data

The varieties we investigate include two ENL varieties, BrE and AmE, three ESL varieties, IndE, SgE and NigE, and three EFL populations, Finnish, French-speaking Belgian and Polish learners. We include two native varieties in response to Hundt and Vogel's (2011) and Deshors' (2017) call for the inclusion of multiple ENL corpora when comparing non-native Englishes; both these studies, among others, find AmE to differ from BrE with regard to progressive usage. As regards the non-native varieties, IndE and SgE were selected for their divergent substrate language backgrounds and evolutionary stage in Schneider's (2007) Dynamic Model – according to the model, IndE is at a less advanced stage of evolution (Phase 3; i.e. begins to develop local standards) compared to SgE (Phase 4; i.e. undergoing a stabilization process). As regards NigE, on the other hand, it is said to be in Phase 3 and already showing some signs of Phase 4 (Schneider 2007: 212), but more importantly, NigE is an African variety of

English which, on the whole, have received less attention than, e.g. Asian varieties. Furthermore, the three ESL varieties chosen situate themselves differently with regard to ENL: SgE has been found to be rather similar to BrE (e.g. Rautionaho 2014), IndE is often characterized by the divergent functions of the progressive (e.g. Mesthrie and Bhatt 2008), thus aligning it further away from ENL on the continuum, and NigE has been shown to differ from BrE in many respects related to the progressive (e.g. higher frequencies, semantically extended uses and higher proportion of present tense usages; Gut and Fuchs 2013). With regard to EFL, we focus on Finnish, Polish and French-speaking Belgian learners. These learner populations were selected based on the fact that Finnish, Polish and French all belong to different language families (that is, the Uralic, Slavic and Romance families, respectively) and none of these languages is Germanic in nature; therefore they are all typologically different from the English language.

As regards the selection of corpora, the ESL varieties, and BrE, are represented by the *International Corpus of English* (ICE) family of corpora (Greenbaum 1991), while the EFL varieties are represented by the *Louvain International Database of Spoken English Interlanguage* (LINDSEI; Gilquin et al. 2010). To represent AmE, we chose the *Santa Barbara Corpus of Spoken American English* (SBCSAE; Du Bois et al. 2000–2005), which is regularly used to substitute for the unfinished spoken part of ICE-US. To ensure compatibility with data drawn from the ICE and LINDSEI corpora, we only included SBCSAE files that represent face-to-face conversation. Similarly, there are differences between the ICE corpora and the LINDSEI corpora that require comment. LINDSEI consists of three different text types: discussion of a set topic, free discussion and picture description. For this study, only the set topic and free discussion sections were used, as the picture description task may induce more frequent use of the progressive form and thus distort the results, and only the students' turns (B-turns) were included in the analysis, i.e., the interviewers' turns (A-turns) were excluded. While there are no text categories in ICE that would be fully comparable to LINDSEI, we limited our focus on Face-to-face conversations (i.e. files S1A-001 to S1A-090). Most earlier studies comparing ESL and EFL have been based on the written sister corpus of the LINDSEI, the *International Corpus of Learner English* (ICLE; Granger et al. 2009), which is relatively well comparable with certain sections in ICE (e.g. student essays and examination scripts, academic writing). Previous studies comparing LINDSEI with ICE are much fewer, and they have resorted to a number of different strategies; Gries and Deshors (2015) use the Class lessons section of ICE, while Götz and Schilk (2011) pair each text type in LINDSEI with a separate text type in

ICE. Having closely examined the LINDSEI discussions, we chose Face-to-face conversations as the most comparable section of ICE, based on similarities in the discussion settings and topics of conversation. We acknowledge the fact that the three corpora are not fully comparable, which may have an effect on the generalizability of our results. However, as long as there are no fully comparable corpora representing spoken[3] ENL, ESL and EFL, we will have to resort to the best possible match we can find, in this case, the Face-to-face conversations.[4]

3.2 Data extraction and annotation of linguistic factors

The progressives were extracted from the corpora using WordSmith Tools Concordancer, with the form *ing preceded by different forms of the auxiliary verb BE (i.e., *am*, *'m*, *are*, *aren'*, *'re*, *is*, *'s*, *isn'*, *was*, *wasn'*, *were*, *weren'*, *be*), with a maximum of five words in between. The non-progressives were extracted (i) manually from SBCSAE and LINDSEI and (ii) automatically[5] from the syntactically annotated (PoS-tagged and parsed) versions of ICE (see e.g. Schneider and Hundt 2012). We did not restrict contexts of use in any way to ensure that all possible contexts where the progressive may occur were included, despite the fact that in Standard English certain linguistic contexts may favor the non-progressive. Note, however, that we did not include perfect progressives, since the combination of a perfect with the progressive is generally rare (Leech et al. 2009: 124). Furthermore, those instances which only superficially resemble progressives, in that they include a form of the verb BE and a present participle, were manually excluded from the data. Such instances include nouns and adjectives, gerunds, appositively used participles, non-finite clauses and the future marker *be going to*. With regard to the non-progressive tokens, we followed the same criteria, where eligible, and additionally, manually excluded a number of constructions that are rare or inexistent in the progressive. These include, for instance, BE *to* V, existential *there*-construction and imperatives (for more detailed information, see Rautionaho and Deshors, forthc.). Furthermore, the study only includes c. 250 progressives and c. 250 non-progressives per variety, randomly chosen from the dataset, and is exclusively based on the lemmas that were observed in both progressive and non-progressive constructions. Table 1 presents an overview of the distribution of progressive and non-progressive constructions within individual investigated English:

Table 1: Overview of the distribution of progressive and non-progressive constructions across English varieties.

Corpus	Progressives	Non-progressives	Total
ICE-GB (BrE)	251	269	520
SBCSAE (AmE)	260	260	520
ICE-IND (IndE)	254	254	508
ICE-SIN (SgE)	252	270	522
ICE-NIG (NigE)	252	251	503
LINDSEI-FI (Fin)	258	257	515
LINDSEI-FR (Fra)	199	304	503
LINDSEI-PL (Pol)	255	255	510
Total	1,981	2,120	4,101

Each token extracted from the data was coded for seven different factors, which are summarized below in Table 2 and some of them briefly discussed below. The coding of all of the factors was cross-checked by each author for the sake of objectivity.

Table 2: Summary of the coding scheme.[6]

Factor	Levels
AKTIONSART	accomplishment, achievement, process, stative
ANIMACY	animate, inanimate
ASPECT	nonprog, prog (dependent variable)
CONTINUUM	EFL, ENL, ESL
SEMANTIC.DOMAIN	activity, aspectual, causative, communication, existence, mental, occurrence
TENSE.MODALITY	modal, past, present
VARIETY	FI, FR, GB, IND, NIG, PL, SG, US
VOICE	active, passive

Aktionsart categories, or situation types (see Vendler 1957; Brinton 1988; Smitterberg 2005), describe situations as made up of different combinations of, minimally, three properties: dynamism (situations consist of either identical or different phases), durativity (situations either last for a period of time or have no duration) and telicity (situations may have a natural end-point after which the

process cannot continue).[7] In the present study, we use the following Aktionsart categorization: States (more or less permanent situations which do not involve dynamism; as in (1)), Processes (dynamic, atelic; as in (2)), Accomplishments (dynamic, telic; as in (3)), and Achievements (dynamic, telic and punctual; as in (4)). Importantly, the Aktionsart category of a token is determined on the basis of the lexical verb and its arguments, as well as by the presence of prepositional phrases or temporal adverbials, as these elements may affect the categorization; a Process becomes an Accomplishment when a countable object (e.g. *eat an apple*) or a PrepP (e.g. *walk to school*) is added. To avoid the Imperfective Paradox,[8] we analyzed the Aktionsart category of a sentence based on the underlying 'unaspectual' form (e.g. [John draw a circle]). Furthermore, it should also be noted that the same lemma may receive a different analysis in different contexts (e.g. THINK may be analyzed as a Process or as a State depending on the context).

(1)   So we have petroleum (ICE-NIG, con_30)

(2)   So your mother looks after them (ICE-SG, S1A-037)

(3)   They are moving out to Portugal next year (ICE-GB, S1A-025)

(4)   She discovers herself her real self (LINDSEI-FR, FR033)

With regard to ANIMACY of the subject, we decided to code the factor in a binary way, i.e. 'animate' vs. 'inanimate', despite the gradient nature of animacy; Strang (1982), for instance, distinguishes between 'human', 'quasi-human or animal' and 'inanimate' subjects. Our binary approach thus regards, for instance, animals as 'animate', and collective nouns such as *family, team* or *company* as 'animate' when they clearly refer to a group of individuals (i.e. human beings). On a number of occasions in our data, the subject is ellipted; in such cases we reconstructed the most likely subject based on evidence from the context, and annotated the animacy of that reconstructed subject. Although animacy of the subject has been discussed as a potential factor in the overall increase of the progressive (see e.g. Hundt 2004; Smitterberg 2005), no clear diachronic trends have been found (see e.g. Kranich 2010: 147). Regionally, however, the preference for the progressive to occur with animate subjects has been shown to vary to some extent in historical BrE, AmE and New Zealand English (Hundt 2004). Regional variation with regard to the animacy of the subject in connection with progressives in Present-Day English has not been, to our knowledge, studied in any detail.

Previous research has indicated the potential importance of morpho-syntactic features in the choice between the progressive and the non-progressive;

Leech et al. (2009: 124) find that, in their diachronic BrE data, present tense progressives increase the most, and that there is also an increase of passive progressives and those that combine with the modal. To model how tense and voice, and the co-occurrence of modal auxiliaries, affect the choice of the progressive over the non-progressive, we annotated each token for 'modal', 'past' or 'present' and 'active' or 'passive'.

3.3   Statistical approach

In order to explore the alternation patterns of progressive and non-progressive constructions across Englishes, we adopted two different statistical techniques. First, we conducted a hierarchical cluster analysis (HAC) to explore degrees of similarity across individual English varieties (i.e. to identify which of our varieties behave similarly with regard to the progressive vs. non-progressive alternation) and then we conducted a logistic regression analysis to examine, within and across individual types of Englishes, how the grammatical contexts of (non-) progressive constructions affect their uses. Our choice to use those two specific techniques reflects our objective to offer an analysis of cross-varietal variation based on the empirical identification of clusters of Englishes rather than the *a priori* theoretically-based assumption that those clusters exist in our data. Besides the general advantage of ensuring that the categorization of our English varieties reflect their behavior in the data, the primary objectives of our cluster analysis are to (i) assess to what degree the clustering of our individual Englishes reflect, empirically, predominant theoretical models of English varieties that distinguish between ENL, ESL and EFL and then (ii) to 'feed' our obtained clusters of Englishes into a logistic regression analysis to ultimately understand how the linguistic contexts of (non-)progressive constructions differ within each of our types of Englishes. Furthermore, an additional benefit of this two-step methodological approach lies in that within the confines of a single analysis we are able to (i) operate at two levels of granularity with regard to English varieties and types (i.e. the more specified level of individual English and the more general level of types of Englishes), (ii) make sure that both levels of granularity are in line with one another and (iii) maximize the power of our regression analysis by predicting speakers' choices of (non-)progressive constructions on the basis of types of Englishes rather than individual varieties. This view is based on the notion that the more levels a variable includes, the less accurate predictions become for each of the levels. So on that basis, types of Englishes (which are smaller in number than individual varieties) will be predicted more accurately compared to the individual varieties. Therefore, in this context, our two statistical techniques should not be approached as independent

of each other. Instead, they should be viewed as connected in that in the regression analysis, the clusters of types of Englishes are used to model the alternation patterns of progressive and non-progressive constructions with a binary logistic regression technique.

Focusing on each of our statistical techniques independently, the HAC approach is an exploratory method which, for the purpose of the present work, provides a way of exploring the cross-varietal similarities and the differences between progressive and non-progressive uses based on a large number of contextual clues. With this method, we are able to compute behavioral profiles of (non-)progressive constructions within each of our Englishes varieties and compare variety-specific profiles across Englishes. As Divjak and Gries (2009: 277) explain, behavioral profiles are comprehensive inventories of elements that co-occur with a word within the confines of a single clause or sentence in actual speech or writing. Statistically, those profiles represent vectors of co-occurrence percentages of a syntactic pattern with all individual predictors' levels (i.e. contextual linguistic features included in the analysis). Behavioral profiles therefore provide form-specific summaries of the semantic and morpho-syntactic behavior of (non-)progressive constructions in each sub-corpus. Based on behavioral profiles, the HAC technique allows us to organize the two constructions in focus by finding dissimilarities between their profiles across investigated English variants and by grouping similar variants together, based on a comprehensive annotation scheme (described in detail in Section 3.2). For the purpose of the present analysis, the individual profiles of (non-)progressive occurrences were computed across the data (i.e., $FR_{.prog}$, $FI_{.prog}$, $PL_{.prog}$, $NIG_{.prog}$, $SG_{.prog}$, $IND_{.prog}$, $GB_{.prog}$, $US_{.prog}$, $FR_{.nonprog}$, $FI_{.nonprog}$, $PL_{.nonprog}$, $NIG_{.nonprog}$, $SG_{.nonprog}$, $IND_{.nonprog}$, $GB_{.nonprog}$, $US_{.nonprog}$) using Gries' (2009) R script Behavioral Profiles 1.01, in relation to the identified semantic and morpho-syntactic predictors. In terms of output, the HAC analysis produces a dendrogram featuring clusters that exhibit high intra-cluster similarity and low inter-cluster similarity and which are, ultimately, all part of a single cluster, the original dataset. In keeping with previous studies (e.g. Divjak and Gries 2006), we chose the Canberra metric as a measure of (dis)similarity and Ward's rule as an amalgamation strategy. For Divjak and Gries (2006: 37), the advantage of the Canberra metric is that it "handles the comparatively large number of zero occurrences of particular features best". The cluster analyses were later validated on the basis of a bootstrap resampling scheme carried out with the R function PVCLUST. Conceptually, resampling consists of sampling repeatedly and randomly, with replacement, from the entire data sample.

In contrast, binary logistic regression is a confirmatory approach that has become relatively customary in many variationist corpus-based studies in World

Englishes research. Broadly, regression analysis is a useful tool that helps us focus on a dependent variable (here ASPECT) and its relation to individual predictors. In the present study, logistic regression modeling helps us identify possible correlations between the predictors and native and non-native speakers' contructional choices within individual types of Englishes (ENL, ESL and EFL). As statistical models, multifactorial regressions are geared towards predicting a particular linguistic item based on a variety of independent co-occurring linguistic features. Therefore, this is an approach with which we can assess how the grammatical context of a linguistic item systematically varies across Englishes as well as the extent of the impact of individual features on the dependent variable. Technically, this approach requires that the data be formatted as a raw annotation table in which all extractions are individually tagged. In terms of data distribution, this approach involves no particular assumption except that the data points are independent of one another. The model we initially used for the regression includes:

- ASPECT as the dependent variable with only two levels: progressive and non-progressive;
- independent variables (specifically CONTINUUM, AKTIONSART, TENSE.MOD and ANIMACY) in the form of main effects;[9]
- all these variables' interactions with CONTINUUM (i.e. including two-way interactions) as additional factors (to see which factors may potentially cause progressive and non-progressive constructions to behave differently in the different types of Englishes).

For the logistic regression, we subjected the dataset to a generalized linear model using the glm (Generalized Linear Model) function in R, using a stepwise modeling approach (with the R function MASS:stepAIC). This ensured the automatic removal from the model of all not-so-useful predictors based on AIC values. In this removal process, not-so-useful interactions were first removed, followed by individual factors that were not significant and did not participate in a significant interaction. To strengthen and validate our results, we then applied a bootstrapping technique (using the calibrate() function in R; see Harrell 2001) to ensure that our predictions were not obtained only when the training and test datasets were identical. This procedure is a cross-validation resampling method that uses a single observation from the original dataset as the validation data and the rest of the observations as the training data. Conceptually, resampling consists of sampling repeatedly and randomly, with replacement, from the entire data sample (Crawley 2007). Before we move on to the results, it should be noted that prior to running the regression model presented in the current paper,

we ran another similar model that included both the CONTINUUM variable (i.e. types of Englishes) and the VARIETY variable (i.e. individual English varieties) as predictors, as opposed to include the former exclusively. Our motivation behind running this more comprehensive model was to assess to what extent the broader category of EFL, ESL and EFL serves as a better or worse predictor of linguistic behavior compared to the more fine-grained classification of individual English varieties. Although that first model yielded a *C* value of 0.81, a likelihood ratio of 1358,7, a *p-value of* $<^{-248}$ and a R^2=0.38, the model also suffered dangerously from extremely high VIF values, many non-attested uses and large confidence intervals which altogether indicated that any conclusions drawn would have been unreliable.

4   Results

Overall, our results are highly interesting, providing support for the fuzzy boundary between ESL and EFL varieties while contrasting with ENL varieties. In what follows, we first present the results of the HAC analysis (Section 4.1) and then we proceed with the more fine-grained multifactorial results (Section 4.2).

4.1   Cluster analysis

The HAC analysis yielded the results presented in Figure 1 which is a dendrogram of the sixteen items, that is, the two constructions in focus across the eight sub-corpora clustered according to their behavioral profile percentages (i.e. based on their co-occurrence patterns with all the independent variables) and the output of the validation of the cluster analysis with PVCLUST for R. The figure shows Approximately Unbiased (AU) as well as Bootstrap Probability (BP) *p*-values. AU *p*-values are computed by multiscale bootstrap resampling, which, according to Suzuki and Shimodaira (2011), is a better approximation to unbiased *p*-values compared to BP values. On that basis, we exclusively focus on AU *p*-values in the remainder of the paper. Broadly, AU values are important indicators of how strongly the data support individual clusters. In the dendrogram, we observe that each cluster is represented by a horizontal line. Importantly, the length of the vertical line is an indication of the distance between clusters and the degree of autonomy of the clustered elements can be inferred based on the length of the vertical lines. So, reading the tree plot from the bottom up, forms clustered early will be more similar than forms clustered late and the longer the line between the clusters, the more autonomous the earlier cluster is from the next cluster it is amalgamated with.
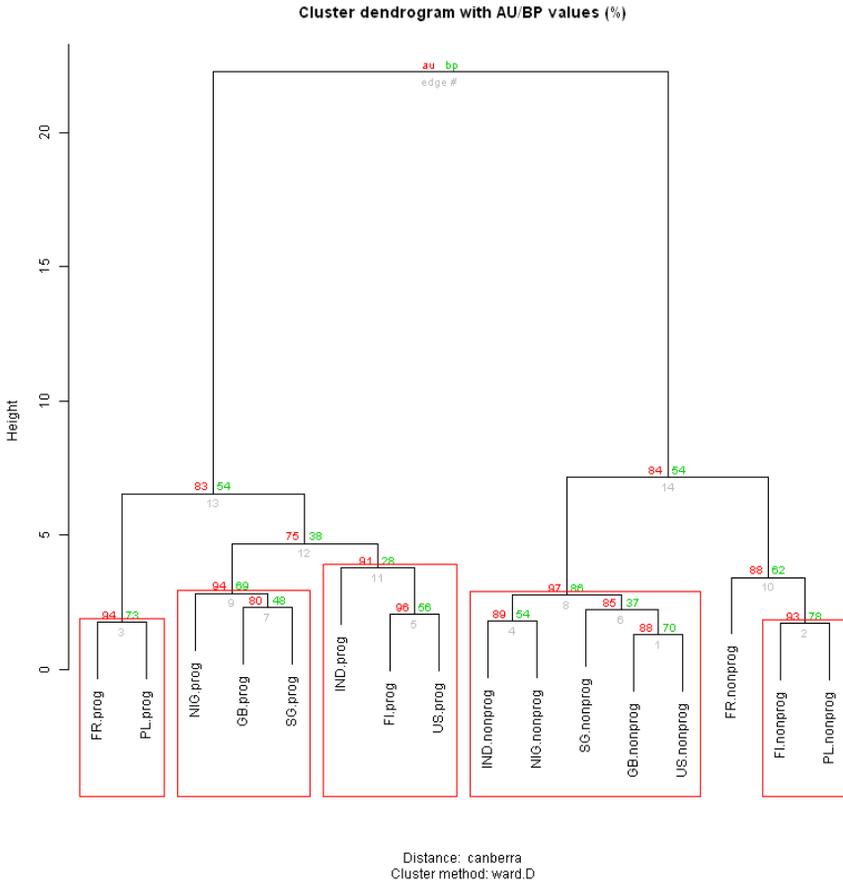
Figure 1: Dendrogram showing the clustering of progressive (prog) and non-progressive (nonprog) constructions in French (FR.$_{prog}$, FR.$_{nonprog}$), Finnish (FI.$_{prog}$, FI.$_{nonprog}$), Polish (PL.$_{prog}$, PL.$_{nonprog}$), Nigerian (NIG$_{prog}$, NIG$_{nonprog}$), Singaporean (SG.$_{prog}$, SG.$_{nonprog}$), Indian (IND.$_{prog}$, IND$_{nonprog}$), British (GB.$_{prog}$, GB.$_{nonprog}$) and American (US.$_{prog}$, US.$_{nonpro}$) English

Based on Figure 1, we can observe the overall degree of (dis)similarity between all the clustered elements. All clusters are amalgamated in one overarching cluster at distance 22 and two main sub-clusters separate at distance 7. As expected, the dendrogram clearly distinguishes between the progressive and the non-progressive constructions. Each construction is clustered independently of the other. The package PVCLUST for R allows us to assess the degree of uncertainty of those clusters and to establish, based on approximately unbiased (AU) *p*-values how strongly the data support the clusters. Generally, the figure shows that the {{IND.nonprog, NIG.nonprog},{SG.nonprog {GB.nonprog, US.nonprog}}} is the most strongly supported by the data with an AU –*p*-value of 100-97=3%. In second place is the {FI.prog, US.prog} sub-cluster (AU *p*-value of 100-96=4%) followed by {NIG.prog, {GB.prog, SG.prog}} (AU *p*-value of 100-94=6%) and {FR.prog, PL.prog} (also with AU *p*-value of 100-94=6%). At this point, it is interesting to note that it is with progressive constructions that we find the highest number of most highly supported clusters: progressive clusters include four clusters with AU values higher than 90 whereas non-progressive clusters include two clusters with AU values higher than 90. Overall, the dendrogram yields highly interesting patterns: while there has been extensive literature on the progressive in non-native Englishes on how the phenomenon is prone to cross-varietal variation (see e.g. Collins 2008; Salles Bernal 2015; Meriläinen et al. 2017), Figure 1 shows how principled this variation is and how dependent on types of Englishes it is. In both main sub-clusters (i.e. the sub-cluster for progressive constructions and the sub-cluster for non-progressive constructions) we observe quite clearly that while, on the one hand, EFLs tend to group together, on the other hand, ESLs and ENLs tend to cluster together. That said, the picture is a little less clear with progressives than it is with non-progressives based on the {IND.prog {FI.prog, US.prog}} cluster towards the middle of the figure, indicating that usage patterns of the progressive are similar in AmE, IndE and Finnish learner English. Furthermore, although the two native Englishes belong to the same cluster, within that cluster, they separate into {NIG.prog {GB.prog, SG.prog}} and {IND.prog {FI.prog, US.prog}}. While this separation is not surprising, given that BrE and AmE have been documented to use progressive marking differently (see e.g. Kranich 2010; Laitinen and Levin 2016), it is nonetheless interesting to note that IndE and Finnish learner English seem to follow the American influence in their usage patterns of progressive constructions, whereas SgE and NigE lean more towards the British influence. In the Finnish context, this tendency to follow American influence may be largely explained by the prominent presence of American TV programs, which are not dubbed but subtitled. It is also interesting to note that the other two EFLs, French and Polish

learner Englishes, which yield the strongest intra-cluster similarities and the strongest inter-cluster differences across all progressive clusters emerge as very similar to one another and less similar to native varieties than the other EFL and ESL varieties. This suggests the possible existence of EFL learner-specific patterns in the uses of progressive marking. Overall, the clustering that emerges from the HAC analysis generally supports a traditional Kachruvian categorization of our Englishes for the regression analysis.

We now move on to the results of the regression analysis to assess which of the linguistic co-occurring features cause the different types of Englishes to vary in their usage patterns of the two alternating constructions.

4.2   Logistic regression

Overall, the regression results show that all significant factors contribute to significant higher-level interactions, i.e. it is only in specific linguistic contexts and only with certain types of Englishes that the factors influence speakers' constructional choice. Therefore, in what follows we will exclusively focus on those interactions and we will not discuss the main effects of individual significant factors. The final regression model reveals a significant correlation between the predictors and speakers' choice of (non-)progressive constructions (Likelihood ratio = 1199.96, $p<^{-234}$), a corresponding relatively strong correlation (R^2=0.38) and a high classification accuracy (76%, C=0.82). Table 3 shows a summary of the significant predictors and interaction terms identified by the model including their coefficients and significance levels. (See Table A1 in the appendix for an overview of the model's confidence intervals.) It should be kept in mind that what is referred to as 'significant' is the contrast between a given factor and its reference level as set for the regression (see Note 9 for a list of all the reference levels in the model). (The codes of statistical significance in Table 3 are to be interpreted as follows: *** stands for $p < 0.00$, ** stands for $0.001 \leq p < 0.01$, * stands for $0.01 \leq p < 0.05$ and '.' stands for $0.05 \leq p < 0.1$.)

Table 3: Overview of the regression model.

	coefficients	Std.Error	Z value	Pr(>|z|)
(Intercept)	1.45107	0.09849	14.733	< 2e-16 ***
CONTINUUMnat_vs_nonnat	-1.13580	0.15647	7.259	3.90e-13 ***
CONTINUUMesl_vs_efl	0.06646	0.22696	0.293	0.769670
AKTIONSARTstate	-2.67963	0.13784	19.440	< 2e-16 ***
AKTIONSARTaccompl	-0.56494	0.16215	-3.484	0.000494 ***
AKTIONSARTachiev	-1.12622	0.28524	-3.948	7.87e-05 ***
TENSE.MODpast	-0.46729	0.14214	-3.288	0.001011 **
TENSE.MODmodal	-3.05660	0.26813	11.400	< 2e-16 ***
ANIMACYinanimate	-0.50582	0.22619	-2.236	0.025338 *
AKTIONSARTstate:TENSE.MODpast	0.59483	0.22307	2.667	0.007663 **
AKTIONSARTaccompl:TENSE.MODpast	-0.86402	0.23133	-3.735	0.000188 ***
AKTIONSARTachiev:TENSE.MODpast	-0.48607	0.40900	-1.188	0.234663
AKTIONSARTstate:TENSE.MODmodal	3.17478	0.40969	7.749	9.24e-15 ***
AKTIONSARTaccompl:TENSE.MODmodal	0.53847	0.41917	1.285	0.198933
AKTIONSARTachiev:TENSE.MODmodal	1.27467	0.75130	1.697	0.089768 .
AKTIONSARTstate:ANIMACYinanimate	-1.11141	0.32805	-3.388	0.000704 ***
AKTIONSARTaccompl:ANIMACYinanimate	0.30952	0.37969	0.815	0.414970
AKTIONSARTachiev:ANIMACYinanimate	0.67152	0.43024	1.561	0.118573
CONTINUUMnat_vs_nonnat:AKTIONSARTstate	0.91859	0.20740	4.429	9.46e-06 ***
CONTINUUMesl_vs_efl:AKTIONSARTstate	0.13454	0.27376	0.491	0.623106
CONTINUUMnat_vs_nonnat:AKTIONSARTaccompl	0.44505	0.22546	1.974	0.048385 *
CONTINUUMesl_vs_efl:AKTIONSARTaccompl	-0.54041	0.29640	-1.823	0.068268 .
CONTINUUMnat_vs_nonnat:AKTIONSARTachiev	0.30804	0.39413	0.782	0.434460
CONTINUUMesl_vs_efl:AKTIONSARTachiev	-0.18527	0.53180	-0.348	0.727554
CONTINUUMnat_vs_nonnat:TENSE.MODpast	0.50059	0.18183	2.753	0.005903 **
CONTINUUMesl_vs_efl:TENSE.MODpast	0.53669	0.25446	2.109	0.034933 *
CONTINUUMnat_vs_nonnat:TENSE.MODmodal	0.39534	0.39079	1.012	0.311716
CONTINUUMesl_vs_efl:TENSE.MODmodal	-0.70755	0.46439	-1.524	0.127602
CONTINUUMnat_vs_nonnat:ANIMACYinanimate	0.79010	0.27553	2.868	0.004137 **
CONTINUUMesl_vs_efl:ANIMACYinanimate	-0.21582	0.36036	-0.599	0.549234

All the predictors included in the regression model (i.e. CONTINUUM, AKTIONSART, TENSE.MOD and ANIMACY) turn out to play a significant part in speakers' decision to use a progressive or a non-progressive construction. Furthermore, the model

shows that those predictors take part in a total of ten different interactions which means that those predictors do not just generally influence constructional alternation patterns but they do so in specific linguistic contexts and only with certain types of Englishes. What is of particular interest at this point and what stresses the importance of considering 'interactions' in our corpus analyses is that based on the model, there is no significant difference between the constructional choices of ESL speakers and those of EFL speakers (*p*=0.76) whereas there is a significant difference between the choices of ENL speakers compared to those of non-native speakers (i.e. EFL and ESL speakers considered together). However, when specific linguistic contexts are considered (i.e. past tense, an Accomplishment situation) the difference between the constructional choices of EFL and ESL speakers becomes significant. Overall, this result calls for a fine-grained degree of analysis when exploring progressive marking in non-native Englishes. In addition to the interactions that involve the English varieties, the model yields interactions between linguistic features *per se* (i.e. AKTIONSART and TENSE.MOD, AKTIONSART and ANIMACY). Those are interactions (or contexts of use) that transcend English varieties. Put differently, in contexts such as those involving a stative situation and an inanimate subject (as in (5)), all three types of English speakers tend to make the same constructional choices. Across the board, we observe that with the predictor AKTIONSART, only States are recurrently favoring non-progressive constructions. This is an interesting result because, while existing literature has devoted much attention to all four Aktionsart categories and their contribution to progressive marking, our results suggest that in the context of understanding what factors influence speakers' decision to select a progressive or a non-progressive, Achievements do not play a strong part (and when they do, they only do so in modal contexts, as in (6)) and Accomplishments only play a significant part in contexts with past tenses (as in (7)).

(5)   It's not as strong as *it looks* (ICE-GB, S1A-042)

(6)   So you *wouldn't find* any individual case fitting very nicely into on one model (ICE-SG, S1A-076)

(7)   I *finished* in nineteen ninety three (ICE-NIG, con_15)

While, with this result, we do not wish to discredit in any way previous work conducted on how Accomplishment, Achievement, Process and State situations all characterize in one way or another the uses of the progressive in World Englishes, our results draw an important distinction between, on the one hand, the contextual features (in this case the specific Aktionsart categories) that we identify as characteristic of the uses of one or the other construction in focus

and, on the other hand, the features that allow us to explain alternating patterns between the two constructions. Put differently, while the co-occurrence patterns of all four Aktionsart categories with a progressive construction may differ and ultimately help us understand better the governing principles behind progressive marking as an isolated linguistic phenomenon, it should not be assumed that all those Aktionsart categories play an influential role in speakers' decision to use or not a progressive construction. In what follows, we discuss our regression results in more detail, starting with the interactions that transcend types of Englishes (Section 4.2.1) and then focusing on those that do distinguish between EFL, ENL and ESL (Section 4.2.2).

4.2.1 Interactions that transcend types of Englishes

According to the regression model, the combined effect of AKTIONSART and TENSE.MOD, on the one hand, and AKTIONSART and ANIMACY, on the other hand, affect speakers' choices of progressive over non-progressive constructions. However, the variation in speakers' constructional choices is not triggered with all Aktionsart categories, tenses or types of subject animacy. In what follows, we specifically focus on the combined effects of the specific Aktionsart categories, tenses and types of animacy that do trigger variation patterns in our data. We first consider the interaction AKTIONSART and TENSE.MOD and then we focus on AKTIONSART and ANIMACY. Figure 2 shows to what extent Achievements and States influence speakers' choices of (non-)progressive constructions in modal contexts. In Figure 2, as well as all the remaining figures, the x indicates the predicted probability, the vertical dashed line marks the mean predicted value for progressive constructions in the data and the error bars show 95 per cent confidence intervals. It should be borne in mind that while the interaction with States is highly significant, the interaction with Achievements is only marginally significant. Thus, observations with regard to the latter can only be tentative at this point. Overall, Figure 2 shows that, expectedly, with States, speakers of all types of English are predicted to opt for non-progressive constructions (as in (8)), due to the inherent incompatibility of stative situations and the durative nature of the progressive construction. The same pattern is observed with Achievements (as in (9)), arising from the fact that Achievements portray situations that include very little or no duration at all, which again is incompatible with the durative nature of the progressive. Building on existing research that already documents the preference of States with non-progressive aspect (Rautionaho and Deshors, forthc.), with our regression results, we are able to pinpoint the exact linguistic context in which this preferential pattern takes place, namely contexts that involve modal uses (as well as, as we will see below in Figure 3, uses with past tenses).

(8)   I mean as a citizen I think I *should know* about it (ICE-IND, S1A-056)

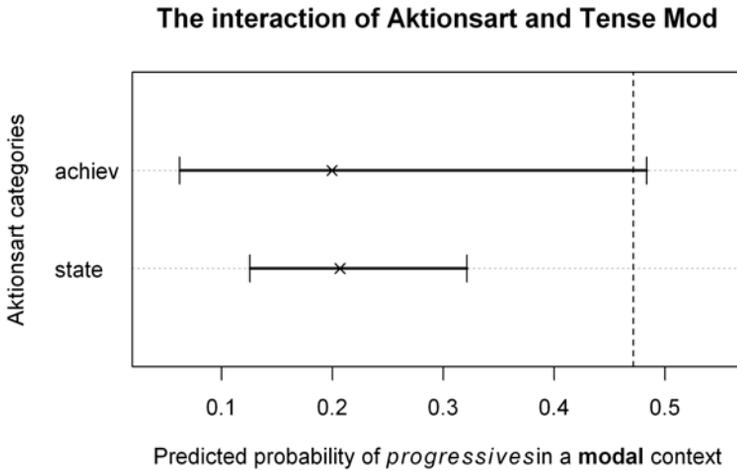(9)   […] we *can't reach* him so we don't know (ICE-NIG, con_36)

The interaction of Aktionsart and Tense Mod

Figure 2: The interaction AKTIONSART:TENSE.MOD (modal)
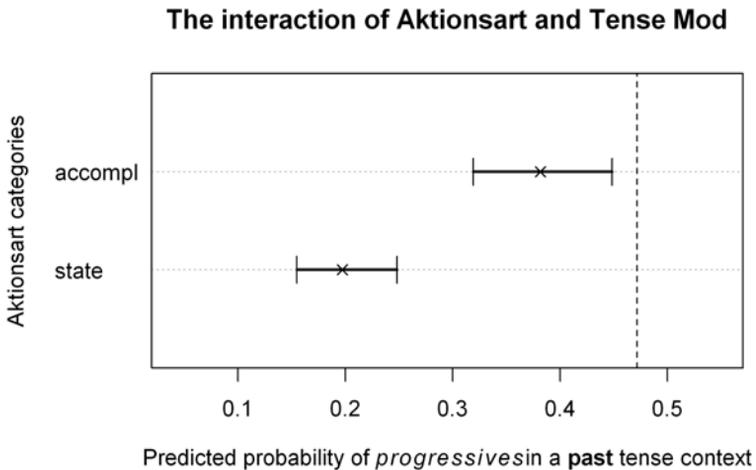
The interaction of Aktionsart and Tense Mod

Figure 3: The interaction AKTIONSART:TENSE.MOD (past)

Continuing with our AKTIONSART:TENSE.MOD interaction (illustrated in Figure 3), let us now focus on contexts of use involving past tenses. In those contexts, and in contrast with modal contexts, we observe that it is Accomplishments and States that significantly influence speakers' constructional choices. Although both aspectual categories show a relatively low association with progressive constructions, they do so in different degrees: Accomplishments attract more progressives than States, which in itself makes sense given that, contrary to States, Accomplishments imply some duration and are thus more compatible with the durative nature of progressives.

Moving on to the next interaction, AKTIONSART:ANIMACY (illustrated in Figure 4), we observe that the influence of ANIMACY over the constructional choices of all our speakers is to some extent very localized. This is mainly based on the observation that speakers are only influenced by this predictor where grammatical subjects are inanimate and where the situation is stative. In fact, it emerges that it is in such context (with inanimate grammatical subjects) that progressive constructions are the least predicted compared to all the other contexts (predictors) included in the study. In other words, the progressive strongly prefers animate subjects, which was to be expected on the basis of earlier diachronic studies, such as Kranich (2010). Furthermore, our results suggest that even though generally States disprefer progressives, this dispreference is not categorical in nature. Depending on the linguistic contexts, the predictability of a progressive construction co-occurring with States can vary to some extent (see section 4.2.2. for more details and examples).
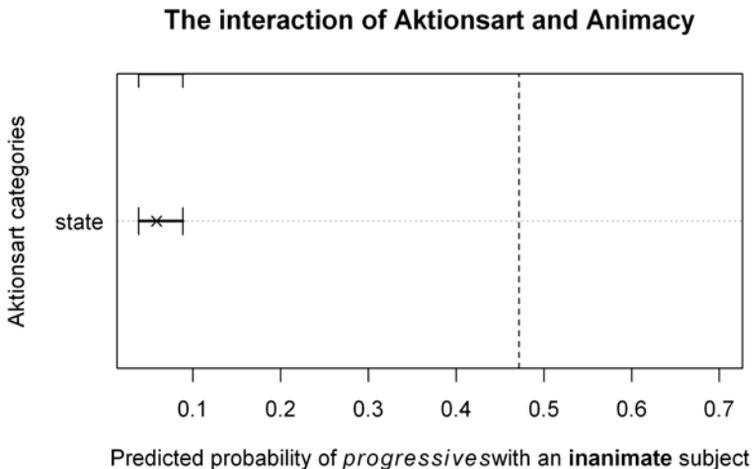
The interaction of Aktionsart and Animacy

Figure 4: The interaction AKTIONSART:ANIMACY

Let us now turn to the other type of interactions, namely those that yield differences across EFL, ENL and ESL. In turn, we will consider how the different types of English speaking populations differ in their constructional choices in contexts that involve different lexical aspect, types of grammatical subject animacy and tenses.

4.2.2 Interactions with the predictor Continuum
Starting with Aktionsart and Accomplishments, specifically, the interaction illustrated in Figure 5 is the only interaction where progressive constructions emerge as more highly predicted than their non-progressive counterpart. More specifically, all three types of English-speaking populations are found to generally prefer progressive marking with Accomplishments. However, it is interesting to note that, despite this general common tendency, individual types of Englishes differ significantly in the degrees to which they prefer the progressive. As our regression model in Table 3 shows, both the differences between, on the one hand, native and non-native speakers and, on the other hand, EFL and ESL speakers are significant. With Figure 5, non-native and native speakers use progressives significantly differently, which is also the case for ESL and EFL speakers. So in contexts where Accomplishments are involved, there is significant variation across the three populations of speakers: ESL speakers are those that are more highly predicted to choose a progressive over a non-progressive (as in (10)). Together, EFL and ESL patterns differ significantly from ENL patterns, with EFL patterns being the least predicted and ESL patterns the most highly predicted. This contrasts with States (illustrated in Figure 6) where only the patterns in native (ENL) and non-native (ESL and EFL together) differ significantly. In other words, whether speakers are EFL or ESL users, their contexts of acquisition and use of English makes no difference on their constructional choices of a progressive or a non-progressive form. With States, the tendency of non-native speakers is to prefer progressive constructions to a lesser degree than native speakers. Another interesting aspect of the result in Figure 6 is that native speakers are the ones who tend to use progressives with States (as in (11); however, the overall preference is for the non-progressive). However, a qualitative analysis of all States in the present data shows that extensions of the progressive into non-delimited States[10] are most frequently found in ESL varieties; thus, while ENL varieties show a higher preference for the progressive to co-occur with States than non-native varieties, it is in ESL varieties, specifically, that we find non-standard use of stative progressives (as in (12)).

(10) They *are developing* curriculum they have asked me to be with them
(ICE-IND, S1A-019)

(11) He*'s being* very fair and dividing his time between us (ICE-GB, S1A-031)

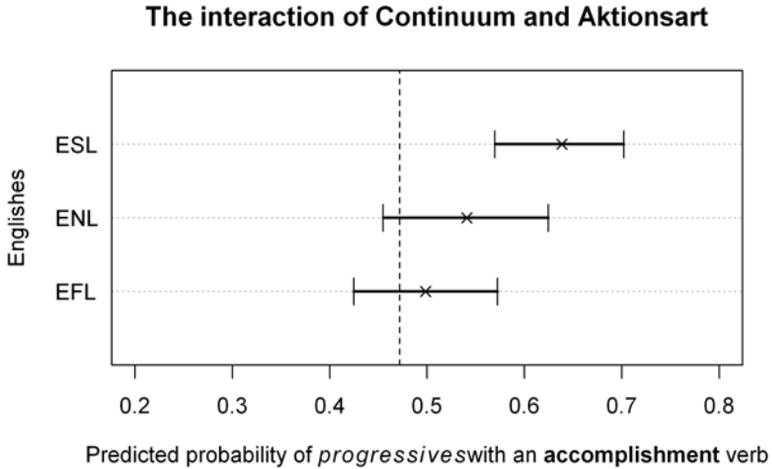(12) But they *must be* uh be *belonging* to some I mean uh good I mean families (ICE-IND, S1A-025)
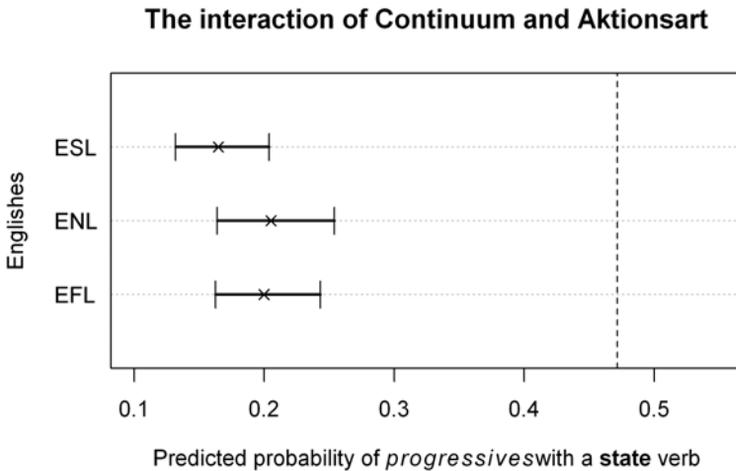
The interaction of Continuum and Aktionsart

Predicted probability of *progressive* with an **accomplishment** verb

Figure 5: The interaction CONTINUUM:AKTIONSART *(Accomplishment)*

The interaction of Continuum and Aktionsart

Predicted probability of *progressive* with a **state** verb

Figure 6: The interaction CONTINUUM:AKTIONSART *(State)*

The interaction of Continuum and Animacy

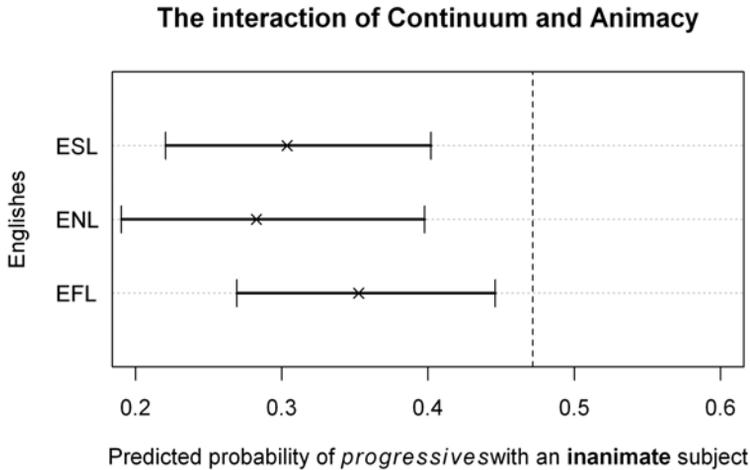Predicted probability of *progressive* with an **inanimate** subject

Figure 7: The interaction CONTINUUM:ANIMACY

Now let us turn to the interaction between CONTINUUM and ANIMACY (illustrated in Figure 7). With this interaction, EFL and ESL users are once again observed to make similar construction choices that, together, differ significantly from those of native speakers. The tendency with this interaction is for native speakers to prefer progressive marking with inanimate grammatical subjects to a significantly lesser degree than EFL and ESL speakers (as in (13)). The fact that non-native speakers actually show a less restricted use of inanimate subjects with the progressive (albeit the preference is, nevertheless, for the non-progressive in all variety types) may indicate that non-native speakers, EFL speakers in particular, are not fully aware of the preference the progressive has for animate subjects, rather than inanimate, or that they are developing different usage norms from those of ENL.

(13) I think that *plan is forming* right now (LINDSEI-FI, FI036)

Finally, let us move on to our last interaction, CONTINUUM and TENSE.MOD (illustrated in Figure 8). With this interaction, we find that in past tense occurrences, all three types of English speakers make constructional choices that significantly vary from one another. More specifically, ENL speakers are the speakers who are predicted to use the highest number of progressive constructions in past tense contexts (that is, more than 50% of the time). With the non-native variet-

ies, we find that EFL users are predicted to opt for progressives more often than ESL users, making EFL speakers' uses of the progressive in past tense contexts more native-like compared to ESL speakers. This tendency of ENL varieties to portray a more or less even distribution of present and past tense progressives and of ESL varieties to prefer the present tense with progressives has been discussed by Rautionaho (2014) and Salles Bernal (2015), among others.[11]
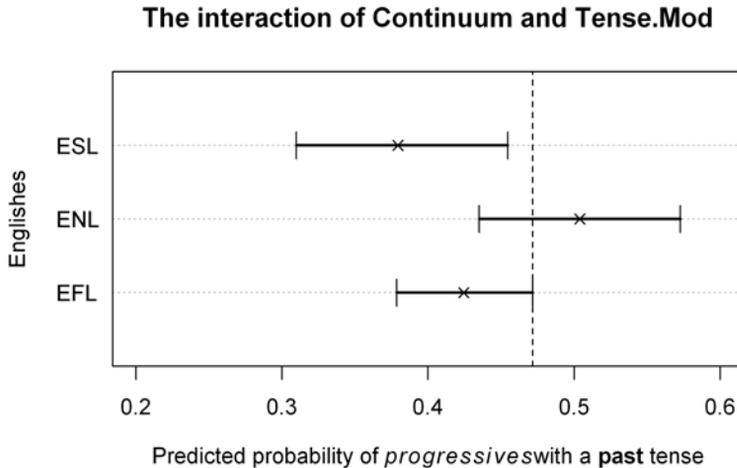
The interaction of Continuum and Tense.Mod

Figure 8: The interaction CONTINUUM:TENSE.MOD

In sum, based on the above results involving Continuum, it seems that each English population has its specific linguistic context in which they favor progressive constructions compared to the other types of Englishes. Specifically, with ESLs, progressives are more highly predicted with Accomplishments, with EFLs, progressives are more highly predicted with inanimate subjects and with ENLs, progressives are more highly predicted with States and past tenses.

5   Discussion

This study set out to investigate the ENL-ESL-EFL continuum from the point of view of the progressive vs. non-progressive alternation. Starting with a cluster analysis and then digging deeper into the data with the logistic regression analysis, our aim was to revisit the currently debated question of whether ENL, ESL and EFL should be approached as dichotomous types of Englishes or whether

they should be seen to form a continuum of native and non-native Englishes. Specifically, we centered our analysis around the following research questions: (i) to what extent do individual English varieties on the ENL-ESL-EFL continuum yield different alternation patterns of progressive vs. non-progressive constructions?, and (ii) which linguistic contextual factors contribute to those different alternation patterns?, and (iii) to what extent do spoken data (as opposed to written data) help us further our understanding of progressive marking in World Englishes?

With regard to (i), to what extent do individual English varieties on the ENL-ESL-EFL continuum yield different alternation patterns of progressive vs. non-progressive constructions?, the cluster analysis showed that the individual varieties do conform well to the traditional Kachruvian model: with non-progressives especially, we observed a clear-cut divide between ENLs, ESLs, and EFLs. With progressives, the result is a little less clear-cut given that although French and Polish learner Englishes form an EFL cluster of their own, Finnish learner English clusters together with IndE and AmE while the two ESL varieties NigE and SgE cluster together with BrE. On the whole, however, the dendrogram supports a dichotomous approach to ENL, ESL and EFL more than it supports the continuum. As regards (ii), which linguistic contextual factors contribute to those different alternation patterns?, the logistic regression analysis draws a relatively complex picture of the (non-)progressive alternation across types of Englishes in that in certain linguistic contexts EFL and ESL speakers use progressives more than native speakers (e.g. with inanimate subjects), in other contexts they use progressives less than natives (e.g. in the past tense) and in yet other contexts (e.g. with Accomplishments) they both differ from one another while being altogether different from ENL. Overall, bringing (i) and (ii) together, while the cluster analysis shows that, at a coarse level of granularity, Kachru's classification makes sense, the regression shows that, when we start digging into the specific linguistic contexts in which progressive and non-progressive constructions are used, the Kachruvian classification does not systematically transpire in the data in a uniform manner. That is to say that drawing conclusions on the validity of the continuum or the dichotomous nature of English types based on cluster results alone may lead to premature generalizations as those conclusions may only hold true for certain linguistic contexts and not others.[12] With regard to progressive marking specifically, this is a crucial point since this is a linguistic feature that has not only been explored alongside a wide range of co-occurring phenomena, including tense, modality, voice, semantic domain of main verbs, subject animacy and functions of the progressive (see e.g. Collins 2008, Edwards 2014, Rautionaho 2014) but it has also fueled much of existing conversations on the validity of the continuum.

In addition to the above, our results bring an interesting perspective on the connection between the type of methodology adopted in studies on the continuum and the position that scholars tend to assume about the dichotomous vs. gradient nature of types of English. Generally, to explore the continuum, scholars have used three types of quantitative approaches: (i) based on (normalized) frequencies of occurrence of progressive constructions (e.g. Hundt and Vogel 2011; Rautionaho 2014; Meriläinen et al. 2017), (ii) based on more sophisticated classification approaches such as cluster analysis and correspondence analysis (e.g. van Rooy 2006; Szmrecsanyi and Kortmann 2011; Edwards and Laporte 2015) and (iii) based on multifactorial logistic regression approaches (e.g. Deshors 2014; Gries and Deshors 2015). What is highly interesting here is that within each type of quantitative approach, all studies tend to come to the same conclusion about the continuum: Hundt and Vogel (2011), Rautionaho (2014), Meriläinen et al. (2017) all support the continuum; van Rooy (2006), Szmrecsanyi and Kortmann (2011) and Edwards and Laporte (2015) all support a dichotomous distinction between EFL, ESL and ENL; and with regressions, there is flexibility with Deshors (2014) supporting the continuum and Gries and Deshors (2015) leaning slightly more towards a dichotomous classification. Specifically, Gries and Deshors (2015: 154) claim that there is no clear-cut position on the continuum discussion: "[W]hile the paradigm gap should be bridged in the sense of analyzing EFL and ESL data together, we should also be open to the possibility that analyzing EFL and ESL together may provide more support for their differences than their similarities". Interestingly, the results we present in the current study follow the above pattern with a cluster analysis that captures distinctions between the three types of Englishes and a regression analysis that allows us to observe that different linguistic contexts cause EFL and ESL to behave more or less similarly and therefore lead more towards a continuum position. The underlying finding here therefore lies in that the conclusions that scholars reach about the validity of the continuum seem to be tightly tied with their choice of quantitative approach. In this context, the present results suggest that, when investigating the continuum through the lens of a particular linguistic item, scholars would benefit from including in their methodological design more than one statistical technique so as to avoid a potentially dangerous methodological bias. So in sum, at a coarse analytical level, our analysis supports the claim that Kachru's classification of English types provides a theoretical framework to explore the linguistic structure of ESL and EFL. At a more fine-grained level, however, we find that only specific linguistic contexts clearly reflect the dichotomy between the three types of Englishes. Ultimately, these results emphasize the importance of not only carefully considering the contexts of use

in which linguistic phenomena are to be assessed but also to bear in mind that, depending on what those contexts consist of, they may lead to different conclusions on the validity of the continuum.

With regard to (iii), to what extent do spoken data (as opposed to written data) help us further our understanding of progressive marking in World Englishes?, based on our results and those presented in Rautionaho and Deshors (forthc.), we are able to identify tentatively three usage patterns of (non-)progressive constructions that seem to be characteristic of spoken language. It should be noted, however, that although Rautionaho and Deshors (forthc.) is, to date, the study which is the most comparable to the present analysis, in that it is a multifactorial analysis of the progressive vs. non-progressive alternation in written language across BrE, AmE, SingE, IndE and Dutch Englishes, the methodological designs of both studies are very similar but not identical. For instance, Rautionaho and Deshors' analysis includes genre as a predictor and is based on individual English varieties rather than types of Englishes. However, despite those differences, we identify the following patterns to further explore in future studies: (i) contrary to written data, the predictor SEMANTIC.DOMAIN does not seem to influence speakers' choice of progressive and non-progressive constructions; (ii) Accomplishment situations and past tense significantly influence speakers' constructional choices but they do not influence writers' constructional choices and (iii) the past tense context has a strong effect with verbs denoting a State and an Accomplishment and distinguishes between, on the one hand, native and non-native constructional choices, and, on the other hand, between EFL and ESL uses. However, in written data, the past tense is not found to influence writers' constructional choices in any of those contexts of use. Ultimately, although these results need to be further investigated, they point towards the need to complement the present work with a multifactorial analysis of progressive vs. non-progressive constructions that will contrast speech and writing as part of a single analysis.

Finally, in addition to the above results, and beyond our three research questions, we would like to briefly return to the initial regression model we computed (see the end of Section 3.3). The main reason behind the poor reliability of the initial model is that there were too few co-occurrences of certain linguistic features within certain varieties, which made it difficult, and in some cases impossible for the model to make predictions on which construction, progressive or non-progressive, a speaker of a particular variety would choose. The fact that the first more comprehensive model was not conclusive is, we believe, an interesting result in itself as it suggests that, at least for the exploration of progressive marking, it may be somewhat misleading to draw conclusions about

types of Englishes, and ultimately the validity of the continuum, solely based on observed patterns at the level of individual Englishes and without assessing quantitatively whether or not those differences can 'safely' lead to realistic conclusions about the continuum. On that basis, it seems that before drawing conclusions, analysts should, whenever possible, complement their results using different statistical techniques.

Notes

1. Even though Deshors (2014) applies both techniques successively in her analysis of the dative alternation in EFL and ESL, her study is set up to explore her data at the level of individual English varieties rather than at the level of types of Englishes.
2. Due to small overall frequencies for aspectual verbs and causative verbs, these two domains were merged for the statistical analysis performed in Rautionaho and Deshors (forthc.). Existence verbs were found to prefer the non-progressive, while Aspect-Causative verbs and Mental verbs showed a preference for the progressive.
3. The situation is slightly better when we want to compare written data: the ICE Student writing sections may be compared to the *International Corpus of Learner English* corpora. Also, the *Corpus of Dutch English* (Edwards 2016) provides data that are, more or less, fully comparable to the entire written part of ICE.
4. We would like to express our gratitude to Sylviane Granger and Sandra Götz for a discussion on the comparability of ICE and LINDSEI.
5. The regular expression used to retrieve non-progressives was '_VB |_VBD|_VBP|_VBZ'. The new versions of ICE have been annotated using the Penn Treebank tagset. The search engine returns the matches in a random order – the first 250 relevant matches were extracted for the purposes of the present study.
6. Although the data were coded for VOICE and SEMANTIC.DOMAIN, those two variables were not included in the regression model due to their highly imbalanced distribution (in the case of VOICE) and, in the case of SEMANTIC.DOMAIN and AKTIONSART, there were too few occurrences of particular AKTIONSART categories co-occurring with specific semantic domains, which would ultimately jeopardize the reliability of the regression model.
7. Some models include more than these three parameters: Vendler's (1957) original classification included the parameter [±Voluntary], while Brinton (1988) added [±Homogeneity] and [±Multiplicity]. For our purposes, the three parameters draw an accurate enough picture.

8.   The Imperfective Paradox (Dowty 1979: 133) refers to the fact that the progressive aspect may affect the Aktionsart categorization value of a situation. In *John was drawing a circle*, the progressive does not entail that the natural end-point was ever reached, i.e. we do not know whether John actually finished drawing the circle.

9.   The reference levels for the independent variables are the following: AKTIONSART = 'process'; TENSE.MODALITY = 'present'; ANIMACY = 'animate'.

10. Following Sharma (2009: 180–181), non-delimited statives (and habituals) are not tied to a specific time-span, i.e. there are no explicit nor implied temporal delimiters (e.g. *these days*) present in the context. Non-delimited statives (and habituals) are thus incompatible with the progressive in StE. For instance, *you must be having waterfalls?* (ICE-IND, S1A-008) refers to a permanent situation, the existence of waterfalls, and is thus a non-delimited instance of a State portrayed by a progressive construction.

11. It should be noted that the preference of the progressive to co-occur with either the present or the past tense is highly context-dependent (see Rautionaho 2014: 104). Depending on the topic of the conversation, speakers choose the more appropriate tense and grammatical aspect; the data representing ENL, ESL and EFL may include different topics of conversation, which, in turn, may affect the result at hand.

12. We acknowledge the fact that Szmrecsanyi and Kortmann (2011) presents a cluster analysis of World Englishes based on a catalogue of over twenty contextual linguistic features. However, despite the demonstrated usefulness of their analysis, by its very nature, their clustering approach does not account for the possible interaction between all the features included in the study.

References

Aarts, Bas, Joanne Close, Geoffrey Leech and Sean Wallis (eds.). 2013. *The verb phrase in English*. Cambridge: Cambridge University Press.

Biewer, Carolin. 2011. Modal auxiliaries in second language varieties of English: A learner's perspective. In J. Mukherjee and M. Hundt (eds.), 7–33.

Bongartz, Christiane M. and Sarah Buschfeld. 2011. English in Cyprus: Second language variety or learner English? In J. Mukherjee and M. Hundt (eds.), 35–54.

Brinton, Laurel J. 1988. *The development of English aspectual systems: Aspectualizers and post-verbal particles*. 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.

Crawley, Michael J. 2007. *The R book*. Chichester: Wiley.

Deshors, Sandra C. 2014. A case for a unified treatment of EFL and ESL: A multifactorial approach. *English World-Wide* 35 (3): 277–305.

Deshors, Sandra. C. 2017. Zooming in on verbs in the progressive: A collostructional and correspondence analysis. *Journal of English Linguistics* 45 (3): 260–290.

Deshors, Sandra C., Sandra Götz and Samantha Laporte. 2016. Linguistic innovations in EFL and ESL: Rethinking the linguistic creativity of non-native English speakers. *International Journal of Learner Corpus Research* 2 (2): 131–150.

Divjak, Dagmar and Stefan Th. Gries. 2006. Ways of trying in Russian: Clustering behavioral profiles. *Corpus Linguistics and Linguistic Theory* 2 (1): 23–60.

Divjak, Dagmar S. and Stefan Th. Gries. 2009. Corpus-based cognitive semantics: A contrastive study of phrasal verbs in English and Russian. In K. Dziwirek and B. Lewandowska-Tomaszczyk (eds.). *Studies in cognitive corpus linguistics*, 273–296. Frankfurt am Main: Peter Lang.

Dowty, David R. 1979. *Word meaning and Montague grammar: The semantics of verbs and times in generative semantics and in Montague's PTQ*. The Netherlands: Kluwer Academic Publishers.

Du Bois, John W., Wallace L. Chafe, Charles Meyer, Sandra A. Thompson, Robert Englebretson and Nii Martey. 2000–2005. *Santa Barbara Corpus of Spoken American English, Parts 1–4*. Philadelphia: Linguistic Data Consortium.

Edwards, Alison. 2014. The progressive aspect in the Netherlands and the ESL/ EFL continuum. *World Englishes* 33 (2): 173–194.

Edwards, Alison. 2016. *English in the Netherlands: Functions, forms and attitudes*. Amsterdam/Philadelphia: John Benjamins.

Edwards, Alison and Samantha Laporte. 2015. Outer and Expanding Circle Englishes: The competing roles of norm orientation and proficiency levels. *English World-Wide* 36 (2): 135–169.

Gilquin, Gaëtanelle. 2015. At the interface of contact linguistics and second language acquisition research: New Englishes and Learner Englishes compared. *English World-Wide* 36 (1): 91–124.

Gilquin, Gaëtanelle, Sylvie De Cock and Sylviane Granger (eds.). 2010. *The Louvain International Database of Spoken English Interlanguage: Handbook and CD-ROM*. Louvain-la-Neuve: Presses universitaires de Louvain.

Gilquin, Gaëtanelle and Sylviane Granger. 2011. From EFL to ESL: Evidence from the International Corpus of Learner English. In J. Mukherjee and M. Hundt (eds.), 55–78.

Götz, Sandra and Marco Schilk. 2011. Formulaic sequences in spoken ENL, ESL and EFL: Focus on British English, Indian English and learner English of advanced German learners. In J. Mukherjee and M. Hundt (eds.), 79–100.

Granger, Sylviane, Estelle Dagneaux, Fanny Meunier and Magali Paquot (eds.). 2009. *The International Corpus of Learner English. Version 2. Handbook and CD-Rom*. Louvain-la-Neuve: Presses Universitaires de Louvain.

Greenbaum, Sydney. 1991. The development of the International Corpus of English. In K. Aijmer, B. Altenberg and J. Svartvik (eds.). *English corpus linguistics: Studies in honour of Jan Svartvik*, 83–91. London: Longman.

Gries, Stefan Th. 2009. *Statistics for linguists with R: A practical introduction*. Berlin, New York: Mouton de Gruyter.

Gries, Stefan Th. and Sandra C. Deshors. 2015. EFL and/vs. ESL? A multi-level regression modeling perspective on bridging the paradigm gap. *International Journal of Learner Corpus Research* 1 (1): 130–159.

Gut, Ulrike and Robert Fuchs. 2013. Progressive aspect in Nigerian English. *Journal of English Linguistics* 41 (3): 243–267.

Harrell, Frank E. 2001. *Regression modeling strategies: With applications to linear models, logistic regression, and survival analysis*. New York: Springer.

Hundt, Marianne. 2004. Animacy, agency and the spread of the progressive in eighteenth- and nineteenth-century English. *English Language and Linguistics* 8 (1): 47–69.

Hundt, Marianne and Joybrato Mukherjee. 2011. Introduction: Bridging a paradigm gap. In J. Mukherjee and M. Hundt (eds.), 1–5.

Hundt, Marianne and Katrin Vogel. 2011. Overuse of the progressive in ESL and learner Englishes: Fact or fiction? In J. Mukherjee and M. Hundt (eds.), 145–165.

Kachru, Braj B. 1985. Standards, codification and sociolinguistic realism: The English language in the outer circle. In R. Quirk and H.G. Widdowson (eds.). *English in the world. Teaching and learning the language and literature*, 11–30. Cambridge: Cambridge University Press.

Kranich, Svenja. 2010. *The progressive in Modern English: A corpus-based study of grammaticalization and related changes*. Amsterdam: Rodopi.

Laitinen, Mikko. (Forthcoming). Placing ELF among the varieties of English: Some observation from typological profiling. In S. C. Deshors (ed.). *Modeling World Englishes: Assessing the interplay of emancipation and globalization of ESL varieties*. Amsterdam/ Philadelphia: John Benjamins.

Laitinen, Mikko and Magnus Levin. 2016. On the globalization of English: Observations of subjective progressives in present-day Englishes. In E. Seoane and C. Suárez-Gómez (eds.). *World Englishes: New theoretical and methodological considerations*, 229–252. Amsterdam/Philadelphia: John Benjamins.

Leech, Geoffrey, Marianne Hundt, Christian Mair and Nicholas Smith. 2009. *Change in contemporary English: A grammatical study*. Cambridge: Cambridge University Press.

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.

Meriläinen, Lea and Heli Paulasto. 2017. Embedded inversion as an angloversal: Evidence from Inner, Outer and Expanding Circle Englishes. In M. Filppula, J. Klemola and D. Sharma (eds.). *The Oxford handbook of World Englishes*, 676–696. Oxford: Oxford University Press.

Meriläinen, Lea, Heli Paulasto and Paula Rautionaho. 2017. Extended uses of the progressive form in Inner, Outer and Expanding Circle Englishes. In M. Filppula, J. Klemola, A. Mauranen and S. Vetchinnikova (eds.). *Changing English: Global and local perspectives*, 191–216. Berlin: De Gruyter Mouton.

Mesthrie, Rajend and Rakesh M. Bhatt. 2008. *World Englishes: The study of new linguistic varieties*. Cambridge: Cambridge University Press.

Mukherjee, Joybrato and Marianne Hundt (eds.). 2011. *Exploring second-language varieties of English and learner Englishes: Bridging the paradigm gap*. Amsterdam/Philadelphia: John Benjamins.

Rautionaho, Paula. 2014. *Variation in the progressive: A corpus-based study into World Englishes.* Tampere: Tampere University Press.

Rautionaho, Paula and Sandra C. Deshors. (Forthcoming.) Progressive or not progressive?: Modeling the constructional choices of EFL and ESL writers. To appear in *International Journal of Learner Corpus Research*, a special issue "Tense and aspect in learner language: Issues and advances in the use of language corpora".

Salles Bernal, Soluna. 2015. Synchronic analysis of the progressive aspect in three varieties of Asian Englishes. *Miscelánea: A Journal of English and American Studies* 51: 87–107.

Schneider, Edgar. 2007. *Postcolonial English: Varieties around the world*. Cambridge: Cambridge University Press.

Schneider, Gerold and Marianne Hundt. 2012. "Off with their heads": Profiling TAM in ICE corpora. In M. Hundt and U. Gut (eds.). *Mapping unity and diversity World-Wide: Corpus-based studies of New Englishes*, 1–34. Amsterdam: John Benjamins Publishing Company.

Sharma, Devyani. 2009. Typological diversity in New Englishes. *English World-Wide* 30 (2): 170–195.

Smitterberg, Erik. 2005. *The progressive in 19th-century English: A process of integration*. Amsterdam: Rodopi.

Strang, Barbara M.H. 1982. Some aspects of the history of the *be + ing* construction. In J. Anderson (ed.). *Language form and linguistic variation: Papers dedicated to Angus McIntosh*, 427–474. Amsterdam: Benjamins.

Suzuki, Ryota and Hidetoshi Shimodaira. 2011. *Hierarchical clustering with P-values via multiscale bootstrap resampling*. https://cran.r-project.org/web/packages/pvclust/pvclust.pdf > (accessed June 27, 2017).

Szmrecsanyi, Benedikt and Bernd Kortmann. 2011. Typological profiling: Learner Englishes versus indigenized L2 varieties of English. In J. Mukherjee and M. Hundt (eds.), 167–187.

Van Rooy, Bertus. 2006. The extension of the progressive aspect in Black South African English. *World Englishes* 25 (1): 37–64.

Van Rooy, Bertus. 2010. Social and linguistic perspectives on variability in World Englishes. *World Englishes* 29 (1): 3–20.

Van Rooy, Bertus. 2011. A principled distinction between error and conventionalized innovation in African Englishes. In J. Mukherjee and M. Hundt (eds.), 189–207.

Van Rooy, Bertus 2014. Progressive aspect and stative verbs in Outer Circle varieties. *World Englishes* 33 (2): 157–172.

Van Rooy, Bertus and Caroline Piotrowska. 2015. The development of an extended time period meaning of the progressive in Black South African English. In P. Collins (ed.). *Grammatical change in English world-wide*, 465–483. Amsterdam: John Benjamins Publishing Company.

Vendler, Zeno. 1957. Verbs and times. *The Philosophical Review* 66(2): 143–160.

Williams, Jessica. 1987. Non-native varieties of English: A special case of language acquisition. *English World-Wide* 8(2): 161–199.

Table A1   Confidence intervals

```
                                                    2.5 %        97.5 %
(Intercept)                                    1.261890674   1.64830517
CONTINUUMnat_vs_nonnat                        -1.444568322  -0.83079602
CONTINUUMesl_vs_efl                           -0.374693917   0.51643953
AKTIONSARTstate                               -2.953567430  -2.41300757
AKTIONSARTaccompl                             -0.881970113  -0.24571426
AKTIONSARTachiev                              -1.683625988  -0.56102514
TENSE.MODpast                                  0.745646544  -0.18804446
TENSE.MODmodal                                 3.607493717  -2.55217902
ANIMACYinanimate                              -0.942542487  -0.05373874
AKTIONSARTstate:TENSE.MODpast                  0.155376405   1.03029350
AKTIONSARTaccompl:TENSE.MODpast               -1.319761183  -0.41255792
AKTIONSARTachiev:TENSE.MODpast                -1.301024830   0.30735767
AKTIONSARTstate:TENSE.MODmodal                 2.360672749   3.97406641
AKTIONSARTaccompl:TENSE.MODmodal              -0.297291848   1.35283525
AKTIONSARTachiev:TENSE.MODmodal               -0.358910672   2.66800181
AKTIONSARTstate:ANIMACYinanimate              -1.773202687  -0.48423362
AKTIONSARTaccompl:ANIMACYinanimate            -0.435271253   1.05652302
AKTIONSARTachiev:ANIMACYinanimate             -0.171616326   1.51888819
CONTINUUMnat_vs_nonnat:AKTIONSARTstate         0.511203195   1.32463976
CONTINUUMesl_vs_efl:AKTIONSARTstate           -0.404541589   0.66948404
CONTINUUMnat_vs_nonnat:AKTIONSARTaccompl       0.003348569   0.88759025
CONTINUUMesl_vs_efl:AKTIONSARTaccompl         -1.123961919   0.03883763
CONTINUUMnat_vs_nonnat:AKTIONSARTachiev       -0.470302808   1.07903234
CONTINUUMesl_vs_efl:AKTIONSARTachiev          -1.228551960   0.86517231
CONTINUUMnat_vs_nonnat:TENSE.MODpast           0.144287883   0.85725364
CONTINUUMesl_vs_efl:TENSE.MODpast              0.038607461   1.03669966
CONTINUUMnat_vs_nonnat:TENSE.MODmodal         -0.389575701   1.15236057
CONTINUUMesl_vs_efl:TENSE.MODmodal            -1.671381033   0.16601640
CONTINUUMnat_vs_nonnat:ANIMACYinanimate        0.246910737   1.32878555
CONTINUUMesl_vs_efl:ANIMACYinanimate          -0.927093799   0.48882641
```