https://doi.org/10.20378/irbo-50209

Valentin Werner

The Present Perfect as a core feature of World Englishes[1]

Abstract: This paper presents central results from a larger corpus-based project (see Werner 2013a; 2013b; 2014) that investigates the usage of the Present Perfect (HAVE + past participle) across World Englishes. It aims at complementing other empirical studies which merely focus on differences between British and American English or which investigate the alternation of the Present Perfect with other time-reference forms.

Findings are based on material from the International Corpus of English (ICE), which has been annotated for various language-internal factors (such as semantics, preceding tense, etc.), so that the distributions and the relative importance of these factors can be analyzed. I employ explorative aggregative methods to find measures of similarity between the various varieties of English under investigation. In addition, this approach allows a systematic investigation of the influence of language-external variables (such as text types, variety types, geographical location) across all varieties. Furthermore, a case study on evidence of an allegedly extended functional range of the Present Perfect in terms of tense-like usage is presented.

The data reveal (i) that the Present Perfect can be seen as a globalized or core feature of world-wide varieties of English, and (ii) that geographical location, variety types, mode of discourse, genres and text types only have a weak effect when associations between varieties are explored; significant groupings across all varieties appear along register lines, however. The case study shows that creative usage in terms of a functional extension of the Present Perfect occurs in the ICE data, albeit largely restricted to informal speech in L1 varieties and to L2 varieties, where influence from both the substrate and through learner language is highly likely. The case study further exemplifies layering between the Present Perfect and its competitor, the Simple Past, in indefinite temporal environments.

Keywords: Present Perfect, Simple Past, involved vs. informational text type, American English (AmE), British English (BrE), International Corpus of English (ICE), L1 varieties, L2 varieties, aggregative analysis, cluster analysis, phylogenetic networks, NeighborNet

1 I would like to thank Ole Schützler and an anonymous reviewer, who both commented on earlier versions of this paper.

Valentin Werner, University of Bamberg

DOI 10.1515/9783110429657-005

1 Introduction

The Present Perfect (henceforth PrPf), both in its standard version HAVE + past participle and in alternative surface constructions with a similar function usage (which will not feature in this analysis),[2] has been among the most frequently discussed issues of English grammar. While earlier research labeled it a "somewhat inconvenient case" (Bauer 1970: 189; see also Labov 1978: 13), 40 years on – and notwithstanding the bulk of literature on the subject (on which see further below and Klein 2009: 54) – the general situation does not seem to have changed, as the PrPf still "eludes a convincing analysis" (Veloudis 2003: 385). I will briefly summarize the main areas of dispute.

A first point that seems difficult is to assign the PrPf to one of the established grammatical classes. To illustrate this point, I will use the categorization approaches as developed in influential descriptive reference grammars of English (see also Werner 2013a). The two relevant approaches as contained in Quirk et al. (1985), Biber et al. (1999), and Huddleston and Pullum (2002) use different kinds of workarounds to achieve an elegant description, which, in turn, leads to inconsistencies in the overall models proposed. For instance, Huddleston and Pullum (2002) label the PrPf a "primary present" and "secondary perfect tense", but state at the same time that "past" and "perfect" form the umbrella category of "past tense" (see Figure 1).

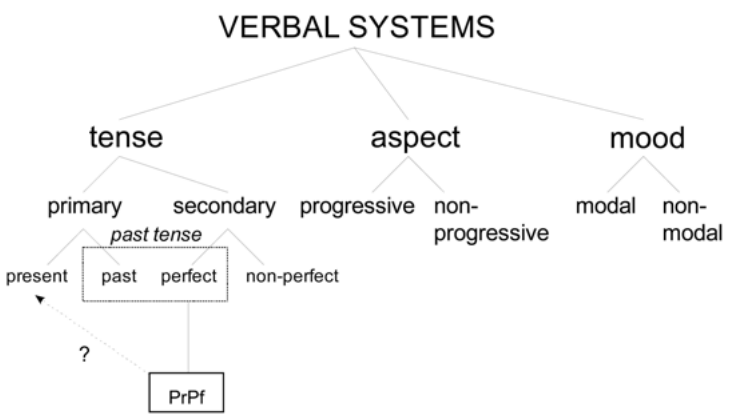

Figure 1: Grammatical categories (= verbal systems) as represented by Huddleston and Pullum (2002); adapted from Werner (2014)

2 Among others, this includes occurrences with auxiliary ellipsis or without any morphological marking at all, BE-perfects, and special cases such as the *after*-perfect (as mainly found in Irish English).

Moreover, some of the grammars listed apply a purely form-based approach. The only reason why the PrPf does not qualify as tense, as is the case in the Quirkian grammar (shown in Figure 2; similarly in Biber et al. 1999), for example, is that tense is taken to be realized morphologically, while aspect is realized syntactically (hence the "morphemic boundary" between the two grammatical classes and the categorization of the PrPf as aspect in this second set of grammars; see further Section 3.2 below).

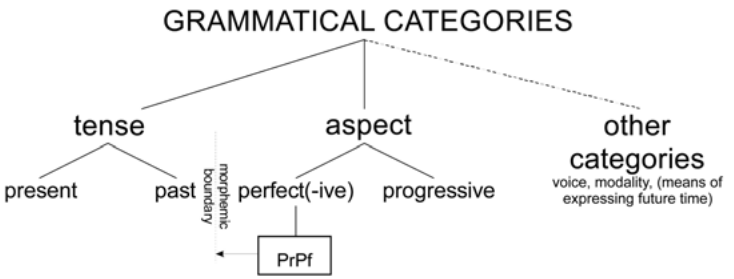

Figure 2: Grammatical categories as represented by Quirk et al. (1985)[3] and Biber et al. (1999); adapted from Werner (2014)

In addition to the analyses in descriptive grammars, a range of theoretical literature on the subject exists, largely revolving around the following topics:

- The PrPf as a grammatical category: tense vs. aspect views vs. further labels, such as "phase" or "status" (see e.g. Bauer 1970; Salkie 1989; Jaszczolt 2009);
- Semantic (and pragmatic) interpretation of the form (see e.g. McCawley 1983; Portner 2003);
- Compositionality of the form (see e.g. Klein 1992; Kortmann 1995).

While all of these works have their merits, some of their inherent weaknesses need to be exposed. First, with regard to semantic readings, depending on the individual author, between one and seven possible interpretations are given and often cross-classifications are possible. Second, another limitation that applies to both descriptive grammars and theoretical treatments of the PrPf alike is that they restrict their focus to a standard variety of British and, less often, American English. Third, many of the models presented are based on introspective analyses and constructed example sentences (but cf. Biber et al. 1999).

3 Means of expressing the future are treated separately in Quirk et al. 1985 (4.41–4.48), while they are included under modals by Biber et al. 1999 (483–497), hence the notation in brackets in the figure.

Starting with Elsness's (1997) pioneering investigation, however, a few corpus-based studies (e.g. Wynne 2000; Schlüter 2002; Hundt and Smith 2009) have overcome the latter difficulty by relying on authentic language data, and have presented quantitative views of the PrPf. Still, while these works have laid the groundwork for the present and other follow-up corpus studies, I would like to argue that they leave some room for improvement. First, some of them apply the semantic models criticized for being based on mere introspection (see above); second, their conclusions partly rely on small and unbalanced corpus material, as has been noted earlier by Schlüter (2006) and Gries (2006), for instance; and third, the majority of the analyses is again restricted to British and American data. Even though the study of World Englishes has been established as a subfield of English linguistics for an extended period now, further varieties (of different types) have been considered in recent years only. Within this area, a distinction between studies focusing on the alternation of the PrPf with other time reference forms, notably the Simple Past (e.g. Davydova 2011, 2016; Yao and Collins 2012; Seoane and Suárez-Gómez 2013; Suárez-Gómez and Seoane 2013; Werner 2013b; see also the contributions in Werner, Seoane and Suárez-Gómez 2016), and within more general "inner life" analyses of the PrPf that consider contextual factors (e.g. Werner 2014) can be drawn. In the present paper, I seek to further extend the empirical perspective on the PrPf within the scope of various World Englishes. Another aspect that the present study will focus on is the systematic inclusion of text type and register – a topic that has been ignored in most analyses as yet to date.

Within the scope of this paper, I will tackle the following issues:

- Is the PrPf a core feature of World Englishes (in terms of overlap between varieties in the distributions of contextual factors) or is nativization (in terms of divergence) observable?
- Along which dimensions (e.g. geographical location, variety types) does variation occur?
- Is there evidence for effects of register, genre or text types within and across varieties?
- Do the data contain evidence of leveling between the PrPf and the Simple Past (SPst) and does this carry implications for the overall grammatical status of the PrPf?

After presenting an outline of the data and the methodology in Section 2, I will show one approach toward relating overlap and divergence of PrPf usage in different varieties in Section 3. This section has two subparts. The first one comprises a global perspective and focuses on text type effects (using aggregative methods). The second part represents a case study. It explores what corpus

data from World Englishes can reveal about the grammatical status of the PrPf (see above) and its potential future development (using a predominantly qualitative approach). I will finish with a general discussion and summary in Section 4.

2 Data and methodology

2.1 The International Corpus of English

In comparison to many of today's mega-corpora of English, such as the *Corpus of Contemporary American English* (450 million words; corpus.byu.edu/coca/) or the even larger *Collins Corpus* (4.5 billion words; www.collins.co.uk/page/The+Collins+Corpus), the scope of the data used for the present study, the *International Corpus of English* (ICE; ice-corpora.net/ice/index.htm), with one million words per regional/national variety, is small. However, it has repeatedly been shown that corpora of this size are well suited for the analysis of grammatical patterns, especially for high-frequency patterns such as the PrPf (Mair 2013: 182; see further Biber 1990). Another motivation for using ICE as a synchronic corpus of different regional L1 and L2 varieties of English is that meaningful comparative analyses across these varieties are only possible with matching corpus components that all adhere to the same compilation principles and thus are as homogeneous as possible with regard to the specification of text categories (see Appendix A), the dating of the data (mostly early 1990s) and the educational background of the informants (adult speakers with at least a completed English-medium secondary school education; Greenbaum 1996: 6; for more detail see Nelson 1996: 28).[4]

For the present analysis, I used the components for Australia (ICE-AUS), Canada (ICE-CAN), Great Britain (ICE-GB), Hong Kong (ICE-HK), India (ICE-IND), Ireland (ICE-IRL), Jamaica (ICE-JA), New Zealand (ICE-NZ), the Philippines (ICE-PHI) and Singapore (ICE-SIN) as well as data from Nigeria (ICE-NIG), East Africa

4 It is evident that in practice there are some differences between corpus components (e.g. as to the time period when the data was sampled or as to individual text types that are used for the individual categories), which – depending on the focus of the research – may influence results. Yet, the ICE family can be viewed as state-of-the-art for the corpus-based linguistic study of World Englishes, as recent additions such as GloWbE (Davies and Fuchs 2015), which can also be used for cross-variety analyses, are restricted to electronically-mediated communication. See the discussion presented in Hundt (2009) for a critical assessment of prospects and limitations of ICE.

(ICE-EA), and the USA (ICE-USA) for the case study.[5] As already indicated, each of the components comprises approximately one million words, while 40% of the material is written, and 60% is spoken. Each component consists of 500 texts of 2,000 words, with many of the 2,000-word text units being composite themselves (see Appendix A).

It is apparent that using this type of data facilitates (i) comparative studies with a focus on the proximity to or distance from a reference variety, most likely British English or American English, (ii) analyses of potential core features across varieties and (iii) accounts of variety-internal variation, which emerges as a growing trend in the study of World Englishes and in variational linguistics in general (Hundt and Vogel 2011: 146; Mukherjee and Schilk 2012: 194).

2.2 Extracting and coding examples

The first step toward a reliable identification of PrPf occurrences in the corpus was to create tagged versions of the corpus files, which are typically available in plain text format only (ICE-GB being one exception). To this end, part-of-speech tags were automatically applied to the corpus data through the *CLAWS* part-of-speech tagger (ucrel.lancs.ac.uk/claws/).[6]

The second step was to create a search string[7] and extract relevant tokens and their context to spreadsheets, which was done with the help of *WordSmith Tools* (Scott 2011). As I retrieved a number of false positives, such as passives, elliptical forms, combinations of modals and PrPfs, and other non-finite forms

5 A note of caution applies, as only the written sections of ICE-USA were available at the time of the analysis. Therefore, additional data, deriving from Yao and Collins (2012) was used (see Section 3.2).

6 The C5 tagset was used for the present study. For the sake of consistency, a plain text version of ICE-GB was created that was subsequently tagged again. Versions of the ICE-components that are tagged with the C7 tagset as well as with semantic tagging are available now (ice-corpora.net/ice/index.htm).

7 The search was for instances of the PP consisting of forms of HAVE (tags *_VHB/*_VHZ) and a past participle (tags *_VBN/*_VDN/*_VHN/*_VVN) within four words to the right. The four-word range was deemed adequate as earlier studies (Schlüter 2002: 103; 2006: 136) found that more than three inserted items between auxiliary and past participle are very rare. The following tag was excluded in the automated search: got_*. The rationale behind this was to exclude the highly frequent combination HAVE *got* (+ NP), as it can virtually always be replaced by a present form HAVE + NP (Wynne 2000: 33); likewise, all instances of HAVE *got to* as a semi-modal that expresses obligation were excluded. In addition, to sharpen the focus of the analysis, progressive forms (tags *_VBG/*_VDG/*_VHG/*_VVG) were excluded due to their almost exclusive association with continuative contexts. The search procedure for the case study will be explicated below (Section 3.2).

(Bowie and Arts 2012), I had to manually exclude them from the analysis (see Werner 2014: 114–117 for a discussion of methodological issues during both the tagging and the identification stage). Still, more than 38,000 data points remained, so I opted for representative random sampling (98% confidence level, 5% margin of error), which eventually left 5,752 data points for the analysis.

The last step in the preparation of the data was to manually code each of the examples according to a number of variables suitable for establishing a kind of grammar of usage (cf. also Biber and Conrad 2009: 216) of the PrPf. The variables had been identified as influencing PrPf usage in other works, and can broadly be categorized into contextual and semantic factors (see Appendix B for underlying models, possible values and examples).

The former group comprises
- presence/absence of temporal adverbials and type of adverbial, as they are commonly viewed as important triggers of the PrPf;
- sentence type, as it is an interacting factor that may determine the presence or absence of temporal adverbials (Schlüter 2002: 242) as well as the semantic interpretation of an instance (Winford 1993: 166);
- preceding time-reference forms, as they may also act as triggers for the PrPf (Davydova 2011: 157).

The latter group consists of
- Aktionsart of the main verb, as inherent lexical aspect carries a substantial part of the meaning of the verb phrase (Schlüter 2006: 143);
- semantic reading of the sentence/clause containing a PrPf occurrence, as a universal property.

2.3 Aggregative analysis

It is one of the principal aims of this study to assess similarity and difference of various varieties on different levels (such as varieties taken as a whole, registers, genres or text types). Therefore, I opted for a multidimensional aggregative statistical approach as an exploratory method used for the identification of latent structure that would not be directly accessible with the help of manual analysis alone. The approach is particularly apt for extended sets of multidimensional data material consisting of a large number of individual data points. Two types are applied: cluster analysis (Romesburg 1984; Manning and Schütze 1999: 495–528) and phylogenetic networks (Huson and Bryant 2006), which have seen a number of applications in linguistics in recent years (e.g. Nichols and Warnow 2008; Szmrecsanyi and Wolk 2011; Kortmann and Wolk 2013; McMahon and

Maguire 2013; Fuchs and Gut 2016; Krug, Schützler and Werner 2016). The main purpose of both types is to graphically represent and reveal relationships of similarity and dissimilarity between different items (varieties, genres and text types in this study). Either unrooted tree-shaped dendrograms or network graphs emerge as the final graphical output, reducing n-dimensional spaces to two-dimensional hierarchical or non-hierarchical representations, respectively (see below).

In the present study, relative values calculated from the absolute values obtained through the coding of the individual factors serve as the input for comparison.[8] These values are entered into comma-separated files (CSV files) to make the data readable for the calculation of the similarity matrices and further processing. In these files, each column represents a register, macro-genre, or text type, and each line contains the values of the same category (e.g. all values of the same category for the factor 'type of adverbial' appear in the same line across all the text types considered).[9] See Table 1 for a snapshot view of such a file.

Table 1: Snapshot view of the top left corner of a CSV spreadsheet used as input for the calculation of the similarity matrix; column labels refer to the ICE text types (of ICE-AUS, in this case); line labels refer to the relative values of the factor 'type of adverbial', adding to 1 (rounded to second decimal place) for each factor

	AUS S1A	AUS S1B	AUS S2A	AUS S2B	AUS ...
Time-position	0.20	0.32	0.26	0.25	...
Span/duration	0.14	0.32	0.42	0.38	...
Frequency	0.34	0.32	0.16	0.12	...
Sequence	0.32	0.04	0.16	0.25	...
...	...	...	...	...	...

These files are subsequently used for aggregative analysis. In the first approach, cluster analysis, items are compared pairwise within a similarity matrix and then fused into clusters that are (depending on the clustering method used)

8 These relative values are the ones presented in Chapter 5 of Werner (2014) for each of the varieties and the respective registers, macro-genres, and text types (unless stated otherwise). All data points had values between 0 and 1. Therefore, no standardization was needed.

9 To calculate statistical measures, *R* was used. For clustering (function *hclust*) I employed the complete linkage method (Manning and Schütze 1999: 505–507) based on the similarity of the two least similar members of a cluster. Rank-based distance matrices were created with the Spearman method. For the creation of the NeighborNet representations I used *SplitsTree* 4.12.3 ("equal angle" method). The most recent version of the program is available at www.splitstree.org (see also Huson and Bryant 2006). As input it requires a nexus file that contains a similarity matrix (see above) created in *R*.

internally maximally similar or minimally dissimilar. In any case, they are highly dissimilar to other clusters and items (Manning and Schütze 1999: 501). In the second approach, data are also clustered, but phylogenetic networks (sometimes also referred to as "phenograms") are used as an alternative (non-hierarchical) means of depicting similarities across varieties. The latter type creates network representations ("NeighborNets") that allow for a more fine-grained analysis. The added value of the NeighborNet representations is that differences, in our case between varieties, and text types, are not reductively shown in terms of absolute cluster membership and categorical (bifurcating) branching. Rather, relative distances to each of the other categories (varieties and text types) that cover the terminal nodes in the graphs are mapped. An interrelation between two nodes is indicated by boxes (splits) in the NeighborNet output. Thus, this method of graphical representation allows us to determine differences between categories that would form members of a single cluster in a hierarchical cluster analysis (see above).

3 Assessing overlap and divergence

3.1 The Present Perfect in the broader perspective

The first section will start from a bird's eye perspective on the data and illustrate the potential of the aggregative methods introduced above. Figure 3 shows a NeighborNet comparing all the data for the eleven varieties under investigation with the individual varieties as nodes or taxa. Colors indicate variety types (L1 = grey; L2 = black, GB as reference variety = underlined). The numbers that are given after each variety indicate the phase label according to Schneider's (2007) dynamic model of postcolonial Englishes. The higher the number in brackets, the further advanced the variety is in the model, so that we could also see the phase labels as some kind of sociolinguistic variety type.

First of all, it emerges that "geolinguistic signals[s]" (Szmrecsanyi 2013: 837) in the data are weak. In other words, geographical location and variety types as external factors only exert a limited influence and clear alignments between varieties of the same type (e.g. in terms of one group of L1 varieties contrasting with L2 varieties or in terms of groups according to the stage labels) do not emerge.

Instead, relationships between the varieties seem to be of a more intricate nature, and a number of further splits deserve closer inspection: As a broad trend, all the L2 (stage 3 and stage 4) varieties apart from Philippine English and Nigerian English are located toward the right hand side of Figure 3, show-

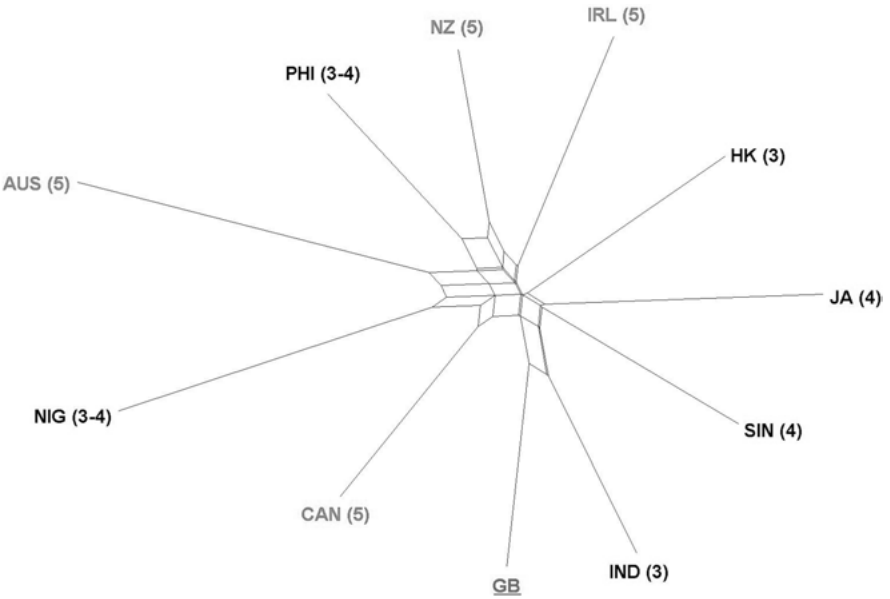

Figure 3: NeighborNet of similarity across ICE components (language-internal factors); numbers in brackets refer to variety type categorization according to Schneider (2007)

ing some association with British English. Another grouping that emerges comprises transplanted L1s (phase 5) varieties (Irish, New Zealand, and Australian English) and Philippine English. The remaining transplanted L1, Canadian English, is also close, but first shares a split with British English, to which it is also close in terms of distance. The L2 varieties align with British English, while the data show that Singapore English and Jamaican English, varieties that have developed further along the evolutionary circle (phase 4), share a split before they merge with the other L2s and British English, the colonial ancestor.

Still, no unambiguous picture emerges. For instance, Indian English as a variety of the L2 type associates closely with British English as its alleged exonormative standard. In contrast, this is not the case to the same extent for Hong Kong English, for instance. In addition, it becomes clear from Figure 3 that Nigerian English covers a special position. Similar to Philippine English, it neighbors the stage 5 varieties Canadian English and Australian English, and, compared to other L2 varieties, appears remote from British English. The latter finding is noteworthy insofar as in Nigeria British English is also still perceived as an exonormative standard to a certain extent (Awonusi 1994: 76; Gut 2012: 2–3).

However, what is most striking is the star-like shape of Figure 3, which indicates that all varieties (with the possible minor exceptions of Australian and Nigerian English) are approximately equidistant from one another. In other words, they differ by approximately the same amount, while sharing many characteristics in terms of quantitative distributions of the factors. Note further that none of the groups, that is, both L1 versus L2 and when the varieties are grouped according to Schneider's (2007) phases, emerges as statistically significant as determined by a bootstrapping test (see also Werner 2014: 305).

When we zoom in, text type effects are evident across varieties, as the cluster dendrogram shown in Figure 4 reveals. This complex dendrogram compares all twelve text type categories (see Appendix A) from the eleven varieties according to the variables described in Section 2.2 above with each other. Each leaf represents one text type, for example the one at the left margin labeled "AUS.w2f" stands for the creative writing category in Australian English.

Above all, note that variety type again does not play a role here, as clusters tend to contain texts from many different varieties. Figure 4 further demonstrates that the distribution of the variables is relatively homogeneous for some of the text types, and this applies mainly to the spoken categories. A case in point is represented by cluster B, which contains 39 out of the 44 leaves representing spoken texts. Given the findings of comparable studies of individual varieties (e.g. Werner 2014), this does not come as a major surprise, as in these analyses, spoken text types regularly emerged as more homogeneous compared to written ones (see also Biber and Conrad 2009: 261). Conversely, clusters A, C and D contain exclusively or predominantly written leaves. Some trends in terms of alignment of individual text types can be established. For instance, seven out of the eleven leaves of the popular writing (w2b) category can be found in B1, six out of eleven of persuasive writing (w2e) cluster in D1 and six out of eleven of creative writing (w2f) in D.

If we change the perspective to the broader ICE macro-genres (see Appendix A), dialogues (19/22) and, even more clearly, monologues (21/22) cluster in B, while almost half of the texts included in non-printed writing agglomerate in clusters C and D (9/22). Printed writing, which comprises a wide range of different text types, is more diverse overall but dominates in B2 (14), C1 (8), D2 (16) and A (6).

It has to be noted that a test for the statistical significance of the clusters in *R* returns only a few clusters on a lower level. No straightforward division (e.g. with a significantly different "spoken" B cluster as indicated above) can be established, which in turn suggests that – with the exception of a few outliers – the data are homogeneous even from a more fine-grained perspective, although some groupings can be observed.

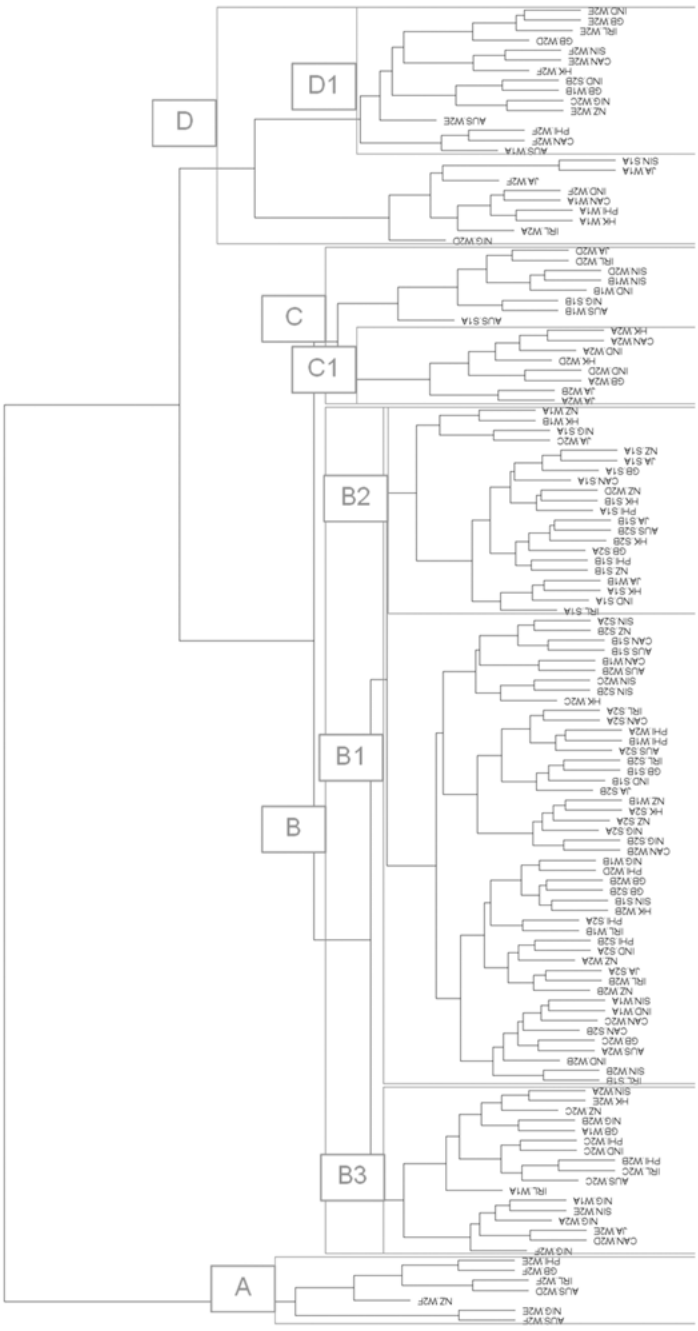

Figure 4: Cluster dendrogram of similarity across ICE text types (cophenetic correlation value = 0.88)

In contrast, significant groupings emerge when we aggregate the data and reorganize them according to the more coarse-grained categories "involved", that is spoken and non-printed, versus "informational" texts, loosely corresponding to the categorization first established in Biber (1988) as "Dimension 1" along which texts or registers may vary. The relevant dendrogram is presented as Figure 5.

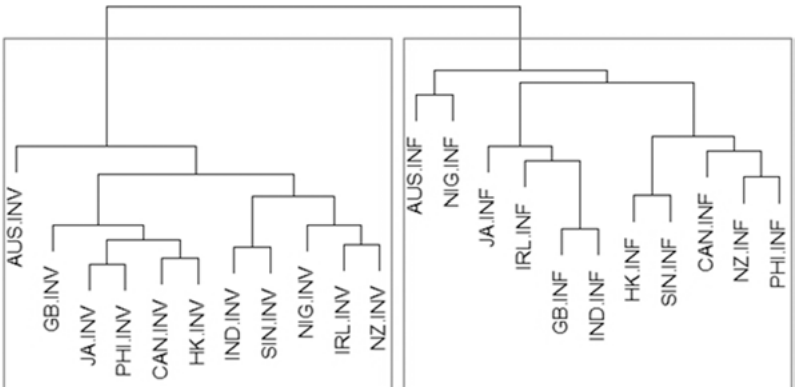

Figure 5: Cluster dendrogram across ICE components: involved (INV) vs. informational (INF) (cophenetic correlation value = 0.76)

It is evident that texts strongly associate with other texts of the same category and that a dichotomy between involved and informational language emerges. All involved leaves (INV) can be found in the left cluster, while the informational leaves (INF) all are found towards the right hand side in Figure 5. A test for statistical significance of the two highlighted clusters confirms the split.

The usefulness of the non-hierarchical perspective as an additional or even alternative means of graphical representation for this type of analysis manifests itself in Figure 6.

The contrast between the two categories involved vs. informational unambiguously appears at one glance in the non-hierarchical network representation. It is illustrated by the larger box-shaped area in the middle of the figure, which separates the two groups and intuitively indicates distance between them. It is also worth noting that, in general, distances between the individual involved nodes are shorter than between the informational nodes, which implies a greater homogeneity of the distributions of the variables in the former type.

In sum, in this aggregated view, we can robustly determine text type associations across varieties. At the same time, this implies that conceptualizations of national varieties as monolithic and clearly separable blocks may be too simplistic (see also Hundt and Vogel 2011), and in particular when individual constructions, such as the PrPf in the present case, and their usage patterns are considered (see Werner 2014: 351–356 for further discussion).

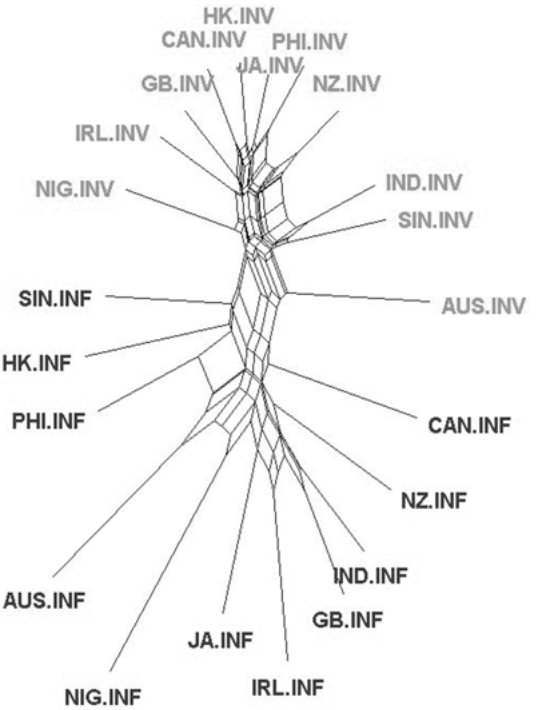

Figure 6: NeighborNet across ICE components: involved (INV) vs. informational (INF)

3.2 Present Perfect versus Simple Past

This section presents a case study that draws its motivation from the well-documented evidence for leveling between the PrPf and the SPst in some varieties of English (eWAVE feature 99; eWAVE feature 100; cf. Lunkenheimer 2012) and in particular discourse types, such as police reports (Ritz 2010) or after-match sports reportage (Walker 2011). It builds on Werner (2013b), now additionally including American English data.

We start from two initial hypotheses. First, in some varieties the PrPf may be used as a narrative tense (in the sense of Quirk et al. 1985), that is, comparable to the SPst, it can be used to create past time reference in combination with definite time adverbials. This represents a development often described in grammaticalization research and is therefore not uncommon from a synchronic typological perspective (Bybee and Dahl 1989: 68–77). Second, we also consider the counter-development, that is, use of the SPst in indefinite temporal contexts typically associated with PrPf use, which is taken as evidence for layering between the two forms (Hundt and Smith 2009: 58; Werner 2013b: 232).

To test these hypotheses, the non-trivial task of identifying contexts where PrPf and SPst are potentially interchangeable with each other without any (fundamental) change in meaning has to be solved. While a certain amount of subjectivity cannot be avoided, I opted for an approach that relies on temporal adverbials as indicators for the contexts under investigation. For the former, I searched for instances where the definite temporal adverbial constraint (as established in Klein 1992) is violated, that is, I searched for combinations of the PrPf with temporal adverbials that typically co-occur with the SPst (x + *ago*, *once*, *yesterday*, *last* + x, in + cardinal number).[10]

Werner (2013b: 229) has shown that in the ICE data relevant examples are scarce and, in terms of register, that they characteristically occur in spoken (and informal) data. While from a quantitative perspective the analysis has suggested a higher salience of tense-like PrPf usage in Asian varieties of English (Indian, Hong Kong, and Philippine English; see also Lunkenheimer 2012: 338–340), no statistically significant differences between the L1 and L2 variety group as a whole have emerged.

However, clear qualitative differences are traceable between variety types. In L1 varieties, occurrences of the PrPf in definite temporal contexts can be explained through pragmatic necessity or as performance errors, which ties in with Rastall (1999: 81–83). One of the types that can be identified comprises examples with iterative statements, such as (1).

(1) For example, the diagnostic and statistics manual (DSM) has been updated twice, *once in 1968, and again in 1980, with a revised version appearing in 1986* (ICE-GB w1a-007)

Here, the definite temporal adverbial constraint seems to be suspended. A second type are examples with afterthoughts or insertions, as (2) or (3).

(2) [...] but we've seen that video *months ago* (ICE-NZ s1b-009)

(3) They have also uhm I think *last year* uh given a list of of six principles (ICE-IRL s2b-001)

10 The view that *in the past* should not be categorized as a temporal adverbial characteristic for SPst contexts, as argued in Werner (2013b: 228), receives further support from the ICE-USA data, as the following examples show: *I think there are some areas we have underemphasized* in the past *that we ought to concentrate on.* (ICE-USA w2b-036); *Because China and the United States have constantly had conflicts* in the past [...] (ICE-USA w1a-012); *In this regard, please ensure that your critical accounting* [...] *analyzes the factors on how the company arrived at material estimates including how the estimates or assumptions have changed* in the past *and are reasonably likely to change in the future.* (ICE-USA w1b-020).

While the hesitation is explicit through *uhm* and *uh* in (3), the corpus annotation of (2) also contains a hesitation marker between *months* and *ago*. It is evident that the violation of the definite temporal adverbial constraint (without further pragmatic factors as described as a prerequisite) is common in colloquial New Zealand English, and the same applies to Australian English (eWAVE feature 100), as illustrated by (4), so these two factors may interact here.

(4) Well he's come on very quickly *last year* (ICE-AUS s2b-017)

Nevertheless, both examples above may be interpreted as instances where the speakers, although they become aware of a "performance error", accept the lack of grammatical well-formedness for pragmatic reasons. Note further that afterthought-like variants may even occur in written texts, as example (5) from the persuasive writing category illustrates.

(5) The Topaz 2 has been used in only two missions, *in 1987.* (ICE-USA w2e-004)

(6) The present tutor training system has been in effect from the inception of the tutoring center *over six years ago*. (ICE-USA w1a-003)

In a similar fashion, (6) exemplifies an instance where the definite temporal information (*six years ago*) is syntactically embedded under an indefinite construction (*from the inception…*) and therefore constitutes some kind of post-hoc specification. A similar interpretation is also conceivable for (1) above, where the focus is first on the indefinite twice and subsequently on definite temporality.

A third issue that seems to play a role is the subjective conception of a situation as recent and relevant by the individual speaker, which is illustrated in (7) with the apparent "recentness" interpretation of the temporal adverbial *this fall*.

(7) Meghan, who will be 4 in March, has also started school *this fall*. (ICE-USA w1b-011)

While premodification, as in (8), is acceptable in both variety types, in L2 varieties structural subjectivity is often conveyed by a combination of definite and indefinite temporal adverbial, as in (9) (see further Werner 2014: 348–349).

(8) Oh I've had some fun *this last week* (ICE-CAN s1a-093)

(9) I would not say that women's issues have *just* started *last year* I would not even say that (ICE-EA s1a-028)

In contrast, the vast majority of the tokens from the L2 varieties exemplify innovative use. On the one hand, motivations behind these uses that clearly violate the definite adverbial constraint (see (10) and (11)) as defined above can be found in the L1 of the speakers (Davydova 2011: 172–173), where comparable structures are acceptable.

(10) You know *yesterday* I have seen some two guys speaking with you (ICE-IND s1a-049)

(11) So he admits Sir what he has stated *yesterday* was was not correct (ICE-IND s2a-063)

In addition to the impact of the substrate(s), learner effects exert some influence. This is mainly the case in terms of a larger variety of grammatical forms for one specific communicative purpose (Werner 2014: 349), which materializes in the present investigation as the occurrence of combinations of PrPfs with definite temporal adverbials, as (12) to (14) illustrate.

(12) Some of them have *once* been my best friends [. . .] (ICE-HK w1b-004)

(13) *Months ago* I have written Sen John Sheffield the head of the US Senate Environment Committee about the problems [. . .] (ICE-PHI s2b-032)

(14) We have mailed you the above DBS Card *2 weeks ago*. (ICE-SIN w1b-019)

In sum, the combination of both of these types of influence, as has previously been argued for other features (e.g. by Schneider 2012: 63–64), seems to provide a plausible motivation for the innovative, tense-like uses of the PrPf in the varieties discussed (see also Werner 2013b: 231–232). The following disclaimer applies, however: Although there is some evidence for creative usage or an extension of the functional range of the PrPf in the ICE data (in particular in the L2 varieties), relevant examples are restricted to informal speech and are rare overall. Although this would not be unlikely from a typological perspective, the data do not support a development of the PrPf into a proper variant of the SPst (or even ousting it). Thus, we can conclude that in the data, the indications for a functional extension of the PrPf are weak at present, as is also shown by converging evidence from other studies (e.g. Elsness 2014: 100). Nevertheless, on a more general note, we may speculate on the potential role substrate and learner effects plays for language change, as changes regularly start out from colloquial usage (Suárez-Gómez and Seoane 2013: 169) and later spread to more formal registers.

In addition to the foregoing qualitative analysis, a brief quantitative view seems worthwhile in order to address the second issue, leveling toward the SPst. While we may not be able to relate all varieties to each other in the same way as shown in Section 3.1 above (due to the lack of comparability of the ICE-USA and ICE-EA data),[11] we can attempt an approximation with the help of assessing the overall "openness" toward the SPst in typical PrPf contexts, again exploiting temporal adverbial contexts, now of the indefinite type. Accordingly, for the identification of examples, a search for adverbials typically associated with the PrPf (*already*, *yet*, *always*, *ever*, *never*, *recently*, *just*, *since*) in both PrPf and SPst contexts was conducted.

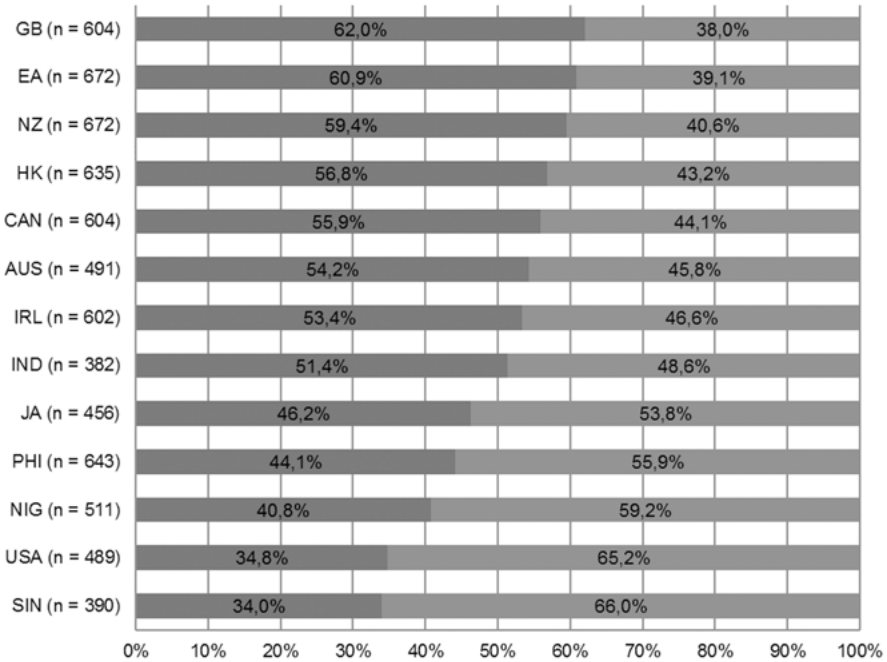

Figure 7: Ranked ratio (relative values) of PrPf (dark grey) vs. SPst (light grey) with indefinite time adverbials

While Werner (2013b) offers an analysis of distributions between SPst and PrPf co-occurring with individual adverbials, the present investigation seeks to arrive at a more global picture. Figure 7 plots the ratio of co-occurrence with either SPst or PrPf averaged across the whole set of indefinite temporal adverbials.

11 As no completed version of ICE-USA is available, the values for the American data are based on the numbers provided in Yao and Collins (2012: 402), who relied on an American corpus that approximates the ICE layout and contains both spoken and written material.

Above all, it establishes that there is considerable variation in the distribution between the two forms. Figure 7 also allows establishing a hierarchy of openness toward the SPst, where Singapore and American English emerge as the most open varieties, while British and East African English appear to be most conservative. This hierarchy largely corresponds to the hierarchy of PrPf friendliness established with a different methodology in Werner (2013b: 213), and further confirms the view of American and British English as two poles on a continuum of openness toward the SPst (see Yao and Collins 2012: 399).

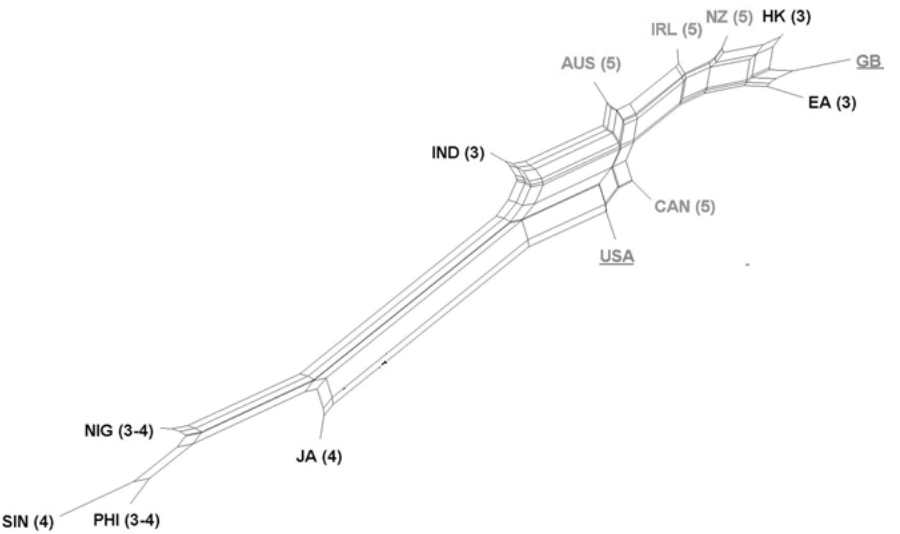

Figure 8: NeighborNet of similarity of SPst vs. PrPf across ICE components; numbers in brackets refer to variety type categorization according to Schneider (2007)

Although American English can be viewed as a variety that is open toward the SPst overall (see Figure 7), variety type plays a part when individual adverbials are considered.[12] We can identify a split along the lines of variety types that is even more clear-cut than the one found in Werner (2013b: 214–216). The NeighborNet shown in Figure 8 reveals a clear distinction between a group of varieties (Singapore, Jamaican, Philippine, and Nigerian English) that emerge as more open toward the SPst (more than 55% of the relevant contexts; see

12 Figure 8 is based on the relative values of the eight indefinite adverbials (see above). As *just*, *recently*, and *since* did not feature in Yao and Collins (2012), the average values across the other 12 varieties are used for American English for these items. This standard procedure applied to missing values in aggregative analyses is further explained in Krug, Schützler and Werner (2016).

above) and a group comprising the remaining varieties. The latter group includes both the reference varieties (American and British English) and the transplanted L1s (phase 5) varieties as well as the nativizing (phase 3) varieties, which arguably still show traces of exonormative orientation and therefore associate closely with the former varieties. In contrast, the group that appears removed exclusively consists of varieties that move toward or have already reached phase 4, which are viewed as being in or entering a process of endonormative stabilization. In spite of the overall distance, American English emerges as the L1 that is closest to the phase 4 varieties. These findings support the usefulness of Schneider's (2007) dynamic model as one approach for the description of World Englishes, although more features would have to be included to assess the overall applicability of this socio-historical approach for the empirical structural description of varieties of English (but see Schneider 2014: 14–15).

4 Conclusion

The results of the study highlight a number of issues. Above all, I hope to have at least partially shown what can be gained from an extension of the research focus beyond the traditional British/American paradigm. While the study of World Englishes has become an established field of English linguistics, there still lies considerable potential in structural analyses that integrate the study of varieties of all kinds. I also want to emphasize again the necessity to systematically include register-, genre- and text type-effects, which can help us to explain variation within and across varieties, as previously noted by Sand (2005: 458) and Schneider and Hundt (2012: 29–30), for example. In this regard, the largely parallel design of the ICE components could be confirmed to be a valuable asset.

The quantitative assessment of the corpus data suggest that the PrPf represents an element of the core grammar of World Englishes. In other words, there was only scarce evidence of the explicit nativization of PrPf usages. Reasons for the convergence may be that grammar represents a linguistic area where convergence is likely in general, while the homogeneous speaker group as represented in ICE and general aspects of globalization may also play a part. To be precise, differences between varieties were of a quantitative rather than a categorical nature, while fine nuances in the distributions and a restricted influence of variety types could be observed. Furthermore, the analyses showed that rather than associations within varieties or varieties of the same type, comparable genres or text types of different varieties emerged as closer to one another. This suggests a revision of the view that conceives of (regional) varieties as monolithic blocks.

The case study provided some evidence for creative language use due to substrate and learner effects, at least as regards individual aspects in the ICE data, and a similar situation applies to further case studies on semantic and pragmatic aspects of alternative surface forms appearing in PrPf contexts (Werner 2014: 322–335; see also Davydova 2011). However, the case for a change of the grammatical status of the PrPf is weak. The second part of the case study exemplified the different characteristic values of layering present in the varieties investigated. In addition, it revealed a fundamental split as to the association of indefinite temporal contexts with either PrPf or SPst along the lines of variety types. For a further assessment in terms of locating varieties in relation to British and American English, it would be desirable to have the full ICE dataset available. In addition, register effects in this domain need to be explored in more detail in the future to obtain a fuller picture.

Closely related to the previous aspect are the final notes on methodology. The present study gave an insight into the benefits of working with corpus data that share a similar layout and largely stick to the same compilation principles. In addition, it demonstrated the usefulness of combining different types of multidimensional aggregative analysis (facilitating the identification of latent structure in big datasets) with further quantitative and qualitative techniques. The extension of aggregative methods to other types of linguistic data (see Krug, Schützler and Werner 2016 for an application on questionnaire data) has the potential to reveal patterns that would otherwise remain hidden.

References

Awonusi, Victor O. 1994. The Americanization of Nigerian English. *World Englishes* 13(1). 75–82.

Bauer, Gero. 1970. The English 'Perfect' reconsidered. *Journal of Linguistics* 6. 189–198.

Biber, Douglas. 1988. *Variation across speech and writing*. Cambridge: Cambridge University Press.

Biber, Douglas. 1990. Methodological issues regarding corpus-based analyses of linguistic variation. *Literary and Linguistic Computing* 5(4). 257–269.

Biber, Douglas, Stig Johansson, Geoffrey Leech, Susan Conrad & Edward Finegan. 1999. *Longman grammar of spoken and written English*. Harlow: Longman.

Biber, Douglas & Susan Conrad. 2009. *Register, genre, and style*. Cambridge: Cambridge University Press.

Bowie, Jill & Bas Aarts. 2012. Change in the English infinitival perfect construction. In Terttu Nevalainen & Elizabeth Closs Traugott (eds.), *The Oxford handbook of the history of English*, 200–210. Oxford: Oxford University Press.

Bybee, Joan L. & Östen Dahl. 1989. The creation of tense and aspect systems in the languages of the world. *Studies in Language* 13(1). 51–103.

Comrie, Bernard. 1976. *Aspect*. Cambridge: Cambridge University Press.

Davies, Mark & Robert Fuchs. 2015. Expanding horizons in the study of World Englishes with the 1.9 billion word Global Web-based English Corpus (GloWbE). *English World-Wide* 36(1). 1–28.

Davydova, Julia. 2011. *The present perfect in non-native Englishes: A corpus-based study of variation*. Berlin: Mouton de Gruyter.

Davydova, Julia. 2016. The present perfect in New Englishes: Common patterns in situations of language contact. In Valentin Werner, Elena Seoane & Cristina Suárez-Gómez (eds.), *Re-assessing the present perfect*, 169–194. Berlin: Mouton de Gruyter.

Elsness, Johan. 1997. *The perfect and the preterite in contemporary and earlier English*. Berlin: Mouton de Gruyter.

Elsness, Johan. 2014. The present perfect and the preterite in late modern and contemporary English: A longitudinal look. In Kristin Davidse, Caroline Gentens, Lobke Ghesquìere & Lieven Vandelanotte (eds.), *Corpus interrogation and grammatical patterns*, 81–103. Amsterdam: Benjamins.

Fuchs, Robert & Ulrike Gut. 2016. Register variation in intensifier usage across Asian Englishes. In Heike Pichler (ed.), *Discourse-pragmatic variation and change: Insights from English*, 183–184. Cambridge: Cambridge University Press.

Greenbaum, Sidney. 1996. Introducing ICE. In Sidney Greenbaum (ed.), *Comparing English worldwide: The International Corpus of English*, 3–12. Oxford: Clarendon Press.

Gries, Stefan Thomas. 2006. Exploring variability within and between corpora: Some methodological considerations. *Corpora* 1(2). 109–151.

Gut, Ulrike. 2012. Towards a codification of Nigerian English: The ICE Nigeria project. *Journal of the Nigeria English Studies Association* 15(1). 1–12.

Huddleston, Rodney & Geoffrey K. Pullum. 2002. *The Cambridge grammar of the English language*. Cambridge: Cambridge University Press.

Hundt, Marianne. 2009. Global English – global corpora: Report on a panel discussion at the 28th ICAME conference. In Antoinette Renouf & Andrew Kehoe (eds.), *Corpus linguistics: Refinements and reassessments*, 451–462. Amsterdam: Rodopi.

Hundt, Marianne & Nicolas Smith. 2009. The present perfect in British and American English: Has there been a change, recently? *ICAME Journal* 33. 45–63.

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: Benjamins.

Huson, Daniel H. & David Bryant. 2006. Application of phylogenetic networks in evolutionary studies. *Molecular Biology and Evolution* 23(2). 254–267.

Jaszczolt, Kasia M. 2009. *Representing time: An essay on temporality as modality*. Oxford: Oxford University Press.

Klein, Wolfgang. 1992. The present perfect puzzle. *Language* 68(3). 525–552.

Klein, Wolfgang. 2009. How time is encoded. In Wolfgang Klein & Ping Li (eds.), *The expression of time*, 39–81. Berlin: Mouton de Gruyter.

Kortmann, Bernd. 1995. Compositionality and the perfect. In Wolfgang Riehle & Hugo Keiper (eds.), *Anglistentag 1994 Graz: Proceedings*, 183–199. Tübingen: Niemeyer.

Kortmann, Bernd & Christoph Wolk. 2013. Morphosyntactic variation in the anglophone world: A global perspective. In Bernd Kortmann & Kerstin Lunkenheimer (eds.), *The Mouton world atlas of variation in English*, 906–936. Berlin: Mouton de Gruyter.

Krug, Manfred, Ole Schützler & Valentin Werner. 2016. Integrating typological profiles and questionnaire data. In Olga Timofeeva, Anne-Christine Gardner, Alpo Honkapohja & Sarah Chevalier (eds.), *New Approaches to English Linguistics: Building Bridges*, 35–66. Amsterdam: Benjamins.

Labov, William. 1978. *The study of NonStandard English*. Urbana: National Council of Teachers of English.

Lunkenheimer, Kerstin. 2012. Tense and aspect. In Raymond Hickey (ed.), *Areal features of the Anglophone world*, 329–353. Berlin: Mouton de Gruyter.

Mair, Christian. 2013. Using 'small' corpora to document ongoing grammatical change. In Manfred Krug & Julia Schlüter (eds.), *Research methods in language variation and change*, 181–194. Cambridge: Cambridge University Press.

Manning, Christopher D. & Hinrich Schütze. 1999. *Foundations of statistical natural language processing*. London: MIT Press.

McCawley, James D. 1983. Tense and time reference in English. In Charles J. Fillmore & D. Terence Langendoen (eds.), *Studies in linguistic semantics*, 96–113. New York: Holt, Rinehart and Winston.

McMahon, April & Warren Maguire. 2013. Computing linguistic distances between varieties. In Manfred Krug & Julia Schlüter (eds.), *Research methods in language variation and change*, 421–432. Cambridge: Cambridge University Press.

Mukherjee, Joybrato & Marco Schilk. 2012. Exploring variation and change in New Englishes: Looking into the International Corpus of English (ICE) and beyond. In Terttu Nevalainen & Elizabeth Closs Traugott (eds.), *The Oxford handbook of the history of English*, 189–199. Oxford: Oxford University Press.

Nelson, Gerald. 1996. The design of the corpus. In Sidney Greenbaum (ed.), *Comparing English worldwide: The International Corpus of English*, 27–35. Oxford: Clarendon Press.

Nichols, Johanna & Tandy Warnow. 2008. Tutorial on computational linguistic phylogeny. *Language and Linguistics Compass* 2(5). 760–820.

Portner, Paul. 2003. The (temporal) semantics and (modal) pragmatics of the perfect. *Linguistics and Philosophy* 26. 459–510.

Quirk, Randolph, Sidney Greenbaum, Geoffrey N. Leech & Jan Svartvik. 1985. *A comprehensive grammar of the English language*. Harlow: Longman.

Rastall, Paul. 1999. Observations on the present perfect in English. *World Englishes* 18(1). 79–83.

Ritz, Marie-Eve A. 2010. The perfect crime? Illicit uses of the present perfect in Australian police media releases. *Journal of Pragmatics* 42. 3400–3417.

Romesburg, H. Charles. 1984. *Cluster analysis for researchers*. Belmont: Lifetime Learning.

Salkie, Raphael. 1989. Perfect and pluperfect: What is the relationship? *Journal of Linguistics* 25. 1–34.

Sand, Andrea. 2005. The effects of language contact on the morpho-syntax of English. In Lilo Moessner & Christa M. Schmidt (eds.), *Proceedings Anglistentag 2004 Aachen*, 449–460. Trier: WVT.

Schlüter, Norbert. 2002. *Present Perfect: Eine korpuslinguistische Analyse des englischen Perfekts mit Vermittlungsvorschlägen für den Sprachunterricht*. Tübingen: Narr.

Schlüter, Norbert. 2006. How reliable are the results? Comparing corpus-based studies of the present perfect. *Zeitschrift für Anglistik und Amerikanistik* 54(2). 135–148.

Schneider, Gerold & Marianne Hundt. 2012. 'Off with their heads': Profiling TAM in ICE corpora. In Marianne Hundt & Ulrike Gut (eds.), *Mapping unity and diversity world-wide: Corpus-based studies of new Englishes*, 1–34. Amsterdam: Benjamins.

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

Schneider, Edgar W. 2012. Exploring the interface between World Englishes and second language acquisition – and implications for English as a lingua franca. *Journal of English as a Lingua Franca* 1(1). 57–91.

Schneider, Edgar W. 2014. New reflections on the evolutionary dynamics of world Englishes. *World Englishes* 33(1). 9–32.

Scott, Mike. 2011. *WordSmith Tools*. Liverpool: Lexical Analysis Software.

Seoane, Elena & Cristina Suárez-Gómez. 2013. The expression of the perfect in South-East Asian Englishes. *English World-Wide* 34(1). 1–25.

Suárez-Gómez, Cristina & Elena Seoane. 2013. They have published a new cultural policy that just come out: Competing forms in spoken and written New Englishes. In Kristin Bech & Gisle Andersen (eds.), *English corpus linguistics: Variation in time, space and genre*, 163–182. Amsterdam: Rodopi.

Szmrecsanyi, Benedikt. 2013. Typological profile: L1 varieties. In Bernd Kortmann & Kerstin Lunkenheimer (eds.), *The Mouton world atlas of variation in English*, 826–843. Berlin: Mouton de Gruyter.

Szmrecsanyi, Benedikt & Christoph Wolk. 2011. Holistic corpus-based dialectology. Brazilian *Journal of Applied Linguistics* 11(2). 561–592.

Veloudis, Ioannis. 2003. Possession and conversation: The case of the category perfect. In Artemis Alexiadou, Monika Rathert & Arnim von Stechow (eds.), *Perfect explorations*, 381–399. Berlin: Mouton de Gruyter.

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

Vendler, Zeno. 1967. *Linguistics in philosophy*. Ithaca: Cornell University Press.

Walker, Jim. 2011. The emergence of the narrative present perfect in British English: Reality or illusion? *Groninger Arbeiten zur germanistischen Linguistik* 53(2). 71–87.

Werner, Valentin. 2013a. The present perfect and definite temporal adverbials: Reference grammars and corpus evidence. *English Language Overseas Perspectives and Enquiries* 10(1). 9–21.

Werner, Valentin. 2013b. Temporal adverbials and the present perfect/past tense alternation. *English World-Wide* 34(2). 202–240.

Werner, Valentin. 2014. *The present perfect in World Englishes: Charting unity and diversity*. Bamberg: University of Bamberg Press.

Werner, Valentin, Elena Seoane & Cristina Suárez-Gómez, (eds.). 2016. *Re-assessing the present perfect*. Berlin: Mouton de Gruyter.

Winford, Donald. 1993. Variability in the use of perfect have in Trinidadian English: A problem of categorical and semantic mismatch. *Language Variation and Change* 5. 141–187.

Wynne, Terence. 2000. *The present perfect: A corpus-based investigation*. Stirling: University of Stirling PhD.

Yao, Xinyue & Peter Collins. 2012. The present perfect in world Englishes. *World Englishes* 31(3). 386–403.

Appendix

Appendix A. ICE text categories (respective number of 2,000-word texts indicated in brackets)

Register/mode of discourse	Macro-genre	Text type category	Text type (detailed)
Spoken (300)	Dialogues (180)	s1a Private (100)	Face-to-face conversations (90) Phone calls (10)
		s1b Public (80)	Classroom lessons (20) Broadcast discussions (20) Broadcast interviews (10) Parliamentary debates (10) Legal cross-examinations (10) Business transactions (10)
	Monologues (120)	s2a Unscripted (70)	Spontaneous commentaries (20) Unscripted speeches (30) Demonstrations (10) Legal presentations (10)
		s2b Scripted (50)	Broadcast News (20) Broadcast Talks (20) Non-broadcast Talks (10)
Written (200)	Non-printed (50)	w1a Student writing (20)	Student essays (10) Exam scripts (10)
		w1b Letters (30)	Social letters (15) Business letters (15)
	Printed (150)	w2a Academic writing (40)	Humanities (10) Social sciences (10) Natural sciences (10) Technology (10)
		w2b Popular writing (40)	Humanities (10) Social sciences (10) Natural sciences (10) Technology (10)
		w2c Reportage (20)	Press news reports (20)
		w2d Instructional writing (20)	Administrative writing (10) Skills/hobbies (10)
		w2e Persuasive writing (10)	Press editorials (10)
		w2f Creative writing (20)	Novels and short stories (20)

Appendix B. Overview of the variables coded

Presence or absence of temporal adverbial

- Type of adverbial (Quirk et al. 1985/Biber et al. 1999)
  - Adverbials of time-position (e.g. *today*, *yesterday*, *afterwards*, etc.)
  - Adverbials of span and duration (e.g. *briefly*, *since* X, *for* X, etc.)
  - Adverbials of frequency (e.g. *daily*, *twice*, *always*, *often*, *never*, etc.)
- Other adverbials of sequence or time relationship between two events (e.g. *already*, *originally*, etc.)
- Aktionsart (Vendler 1957; 1967/Comrie 1976)
  - Activity verbs (e.g. *run*, *eat*, *fly*, etc.)
  - Accomplishment verbs (e.g. *build*, *draw (a circle)*, *run (a mile)*, etc.)
  - Achievement verbs (e.g. *discover*, *reach (the top)*, *cross (the river)*, etc.)
  - State verbs (e.g. *love*, *hate*, *be*, etc.)
- Sentence type
  - Positive statements (e.g. *I've been in this limbo for too long*; ICE-CAN s1a-037)
  - Negative statements (e.g. *Their performance has not been good*; ICE-NZ w2c-014)
  - Questions (e.g. *Are there any other things you have left out*; ICE-JA s1b-078)
- Semantics (Schlüter 2002)
  - Indefinite past – single acts/events (e.g. *the teacher who taught us Bridge has gone*; ICE-HK s1a-042)
  - Indefinite past – multiple acts/events (e.g. *I've probably seen them a dozen times*; ICE-AUS s1a-071)
  - Continuative past – continuous acts/events (e.g. *Complaints have been known to be made when such situations have occurred, this is why the rule now rigidly applies*; ICE-IRL w2d-017)
  - Continuative past – states (e.g. *The economic team of the present administration during its first year has been superior in cohesiveness effectiveness and clarity in purpose*; ICE-PHI s2b-026)
- Preceding time reference