



**University of
Zurich**^{UZH}

**Zurich Open Repository and
Archive**

University of Zurich
University Library
Strickhofstrasse 39
CH-8057 Zurich
www.zora.uzh.ch

Year: 2018

It is time that this (should) be studied across a broader range of Englishes: a global trip around mandative subjunctives

Hundt, Marianne

DOI: <https://doi.org/10.1075/veaw.g61.09hun>

Posted at the Zurich Open Repository and Archive, University of Zurich

ZORA URL: <https://doi.org/10.5167/uzh-151539>

Book Section

Accepted Version

Originally published at:

Hundt, Marianne (2018). It is time that this (should) be studied across a broader range of Englishes: a global trip around mandative subjunctives. In: Deshors, Sandra C. Modeling World Englishes. Assessing the interplay of emancipation and globalization of ESL varieties. Amsterdam: John Benjamins Publishing, 217-244.

DOI: <https://doi.org/10.1075/veaw.g61.09hun>

It is time that this (*should*) be studied across a broader range of Englishes: A global trip around mandative subjunctives*¹

Marianne Hundt, Zurich University

Abstract

English has a choice between modal verb *should* and a subjunctive in subordinate clauses following mandative expressions such as *recommend*, *request* or *require*. Previous research found that the subjunctive has seen a revival in the twentieth-century in a change led by American English with British English and settler varieties lagging behind. Studies of the subjunctive in second-language varieties of English (ESL) are scarce, and typically look at only one ESL variety, comparing its text frequency with that observed in first language varieties. Previous research also looked at the distribution of subjunctives across spoken and written registers, their co-occurrence with active and passive voice, and/or with individual triggers, but these factors have not yet been studied as predictor variables for the choice between a subjunctive and a modal construction. On the basis of the *International Corpus of English*, this paper investigates the choice between mandative subjunctives and modal periphrastic constructions with *should* across a broad range of World Englishes with a view to modelling the relative strength of external predictor variables such as ‘variety’ and ‘medium/register’ as well as internal factors like ‘lexical trigger’ and ‘verb’. It uses evidence from the *Global Web-based English* corpus for a follow-up study on the importance of ‘lexical trigger’ on a subset of the varieties, since ICE corpora are too small to provide robust evidence on this. The findings do not lend themselves to straightforward interpretation within an individual model of World Englishes.

Keywords: mandative sentences, subjunctive:modal alternation, probabilistic grammar approach

1. Introduction

The subjunctive in English has often been described as a ‘moribund’ (Fowler 1926) or near-extinct grammatical category (e.g. Harsh 1968). This view is occasionally maintained up until the end of the twentieth century; Denison (1998: 263) maintained that “[i]n Br[itish] E[nglish] the present subjunctive [...] has retreated to high-flown literary or legal language [...]”. Yet, in subordinate clauses following mandative verbs, nouns and adjectives such as *demand* or *important* the subjunctive form has been increasing again in the twentieth century (e.g. Övergaard 1995, Leech et al. 2009).² The mandative phrase *it is time* from the

*This paper is dedicated to Christian Mair on the occasion of his 60th birthday.

¹ I would like to thank Nina Benisowitsch and Carlos Hartmann for help with data extraction at various stages of the project.

² A notable exception to this view is a recent study by Kastronic and Poplack (2014), who combine data from corpora of Early and Late Modern (British) English with a corpus of Present-Day Canadian English and, using a very different methodology from previous studies, argue that the mandative

title can be used to illustrate the variable context, i.e. the choice between a subjunctive form (typically distinct after third person singular subjects, as in example (1)), a periphrastic construction with a modal verb (typically *should*, as in (2)) or, in some varieties, an indicative, as in example (3):³

- (1) The rights and responsibilities of marriage are civil, they are legal, and now *it is time that they be made* equal. (GloWbE, US, General)
- (2) This is the English practice and, if it is not yet definitely settled in this country, *it is time that it should be?* (GloWbE, IE, General)
- (3) I recognise *it is time that* the unabated issue of dangerous dogs *is grabbed* by the scruff of the neck and brought firmly under control. (GloWbE, GB, General)

Previous research has shown that the revival of the mandative subjunctive can be observed in varieties of English as a first (ENL) or institutionalized second (ESL) language, and that this recent/ongoing change is spearheaded by American English (AmE) (e.g. Övergaard 1995, Hundt 1998). Various corpus-based studies have looked into the diachronic spread of subjunctive and regional variation, but the range of ESL varieties in previous research has typically been restricted to only one (e.g. Sayder 1989, Schneider 2005, 2011) or a few selected varieties (e.g. Peters 2009). Moreover, no previous research has looked into the relative importance that contextual and linguistic factors play in the choice between a subjunctive and the periphrastic construction with modal *should*. It is the aim of this chapter to address this research gap using data from the *International Corpus of English* (ICE) and the corpus of *Global Web-based English* (GloWbE).

In the following, I will briefly summarize the most important findings from previous corpus-based research into mandative subjunctives (section 2) and give a detailed definition of the variable context, data retrieval and background on the factors included in the analysis (section 3). In addition to summary statistics, I will combine random forest analysis and conditional inference trees to probe into the relative importance that contextual and linguistic factors play in the choice between subjunctive and modal periphrasis (section 4). Which model of English as a World Englishes might best explain the patterns of variation found in the corpus data will be the subject of the concluding discussion (section 5). As Mair

subjunctive has not increased over the past few hundred years. This discrepancy is largely due to a very different way of defining the variable context (see section 3).

³ The choice can also be avoided in various ways, e.g. by resorting to a *for-to*-infinitival construction (*They asked for Mr Robinson to resign*) or a nominal instead of a clausal complement (*They asked for Mr Robinson's resignation*). These options are not included as variants in this paper.

(2017: 105) points out, while models of World Englishes are not usually developed with a view to corpus-based research, they “can be seen as presenting a blueprint for a corpus-linguistic research agenda in the field.”

2. Previous research: Recent and ongoing change in World Englishes

One of the earliest studies, based on data from the Brown and LOB corpora and a set of pre-defined mandative triggers, found that the mandative subjunctive was used more often in American than in British English (Johansson & Norheim, 1988). Follow-up studies using the same set of triggers and evidence from the Brown family of corpora (Hundt 1998, Hundt & Gardner 2017, Waller 2017) show that (a) the significant increase in AmE took place in the early years of the twentieth century and (b) the mandative subjunctive has been gaining ground in British English (BrE) in the second half of the twentieth century. However, the change in BrE happens at a much slower rate than in AmE and might already be levelling off at a lower level, judging from the web-based BE06 evidence (see Figure 1).⁴

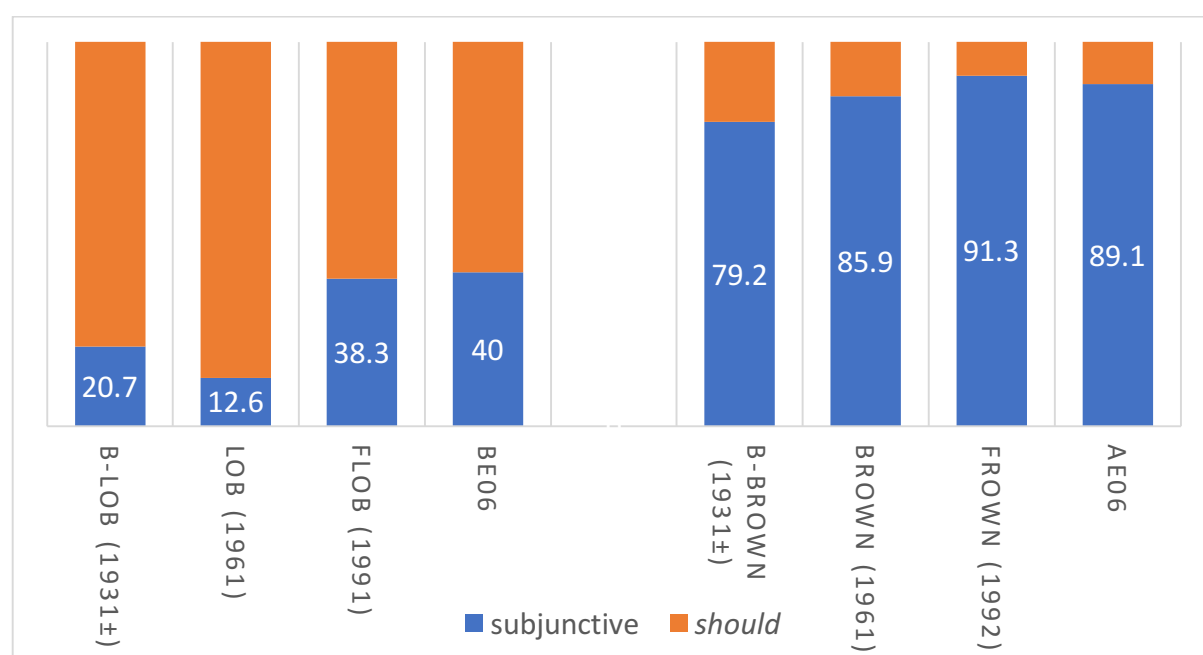


Figure 1: Mandative subjunctives (%) in the Brown-family corpora; variable contexts: B-LOB (1931[±]) = 92; LOB (1961) = 111; FLOB (1991) = 117; BE06 = 70; B-Brown (1931[±]) = 96; Brown (1961) = 134; Frown (1992) = 115; AE06 = 92)

⁴ The data for LOB and Brown are from Johansson & Norheim (1988), those for FLOB and Frown from Hundt (1998); data for the extended Brown family (for the 1930s and 2006, respectively) come from Hundt & Gardner, (2017) and Waller 2017. Note that Waller (2017) finds that proportions of variants in mandative sentences are based on slightly different definitions of the variable across studies, for instance with respect to the inclusion/exclusion of bare forms with plural subjects after past tense triggers (see example (8) below). The overall development and general regional variation are not affected by these differences.

Corpus data also provides evidence that settler varieties such as Australian (AusE) and New Zealand English (NZE) (e.g. Hundt 1998, Peters 1998; Collins 2015) occupy an intermediate stage between BrE and AmE, with a slight preference of subjunctives over periphrastic constructions. Ongoing change in ENL varieties has typically been explained in terms of American influence/Americanization (see e.g. Övergaard 1995: 89 or Mair 2006: 193).

For ESL varieties, previous research based on Brown-type corpora found that they either used similarly low proportions of mandative subjunctives as BrE (e.g. Sayder 1989 or Schneider 2000 on Indian English) or that they aligned with AmE (e.g. Schneider 2005, 2011 and Collins et al. 2014 on Philippine English). Typically, studies on ESL varieties compare only one of these with ENL reference varieties. A notable exception is Peters (2009), who compares evidence on Philippine English (PhilE), Indian English (IndE) and Singapore English (SingE) with three ENL varieties (BrE, NZE and AusE). Her study, which uses six lexical triggers⁵ and both written and spoken evidence from the respective components of ICE confirms previous results for PhilE and IndE, i.e. the affinity with the respective matrillect.⁶ However, SingE turns out to prefer subjunctives over modal constructions, thus casting some doubt on the historical lineage as the best explanation for the usage patterns in ESL varieties. Moreover, for IndE, her study reveals regular use of deontic and quasi-modals in subordinate clauses after mandative triggers, i.e. variants not typically included in the study of the variable context; she also finds evidence of nativized patterns that avoid the choice between a subjunctive and the modal construction in ways different from ENL varieties (Peters, 2009: 130).

With respect to variation across speech and writing, earlier studies have shown that the subjunctive is not limited to formal, written language. On the contrary, corpus data have consistently shown that the subjunctive is regularly attested in speech (e.g. Hoffmann 1997, Hundt 1998, Schneider 2005), though how this factor may interact with regional variety has not yet been discussed on the basis of conclusive evidence (see Peters 2009: 130).

⁵ She searched the ICE corpora for variant forms of the verbs *demand*, *move*, *recommend*, *request*, *require* and *suggest* as well as their related nouns (Peters, 2009: 133).

⁶ Note that Sedlatschek (2009: 286-88), on the basis of a small set of verbs invested in internet-based data, finds IndE to be more conservative than its matrillect.

This short review of previous research shows that additional investigation on the mandative subjunctive in World Englishes is warranted for various reasons. First, comparison across different regional varieties is confounded by the fact that previous research is not necessarily based on the same set of mandative triggers (see also section 3). In addition, studies typically focussed on frequency differences and/or historical developments but rarely looked at factors that might influence the choice between a subjunctive and the periphrastic variant. In other words, there is very little research across ENL and ESL varieties on micro-variation in the use of mandative subjunctives in relation to register, mandative trigger, negation, etc. and none on the relative importance that such factors may have. Is register, for instance, a more important factor than regional variation, as in various other studies on grammatical variation across ENL and ESL varieties? Does the verb *be*, the most clearly subjunctive form after mandative triggers, provide a stronghold for the subjunctive across all World Englishes, as previous research suggests (e.g. Turner 1980: 275)? How frequent are negated subjunctives and are they typical of formal, written language? Is *that*-omission in mandative sentences associated with a particular region and/or register? This chapter uses evidence from a broad range of ENL and ESL varieties to answer these questions against the backdrop of existing models of World Englishes.

3. Data and methodology

3.1 Corpus data of World Englishes

This paper uses data from a total of 10 ICE components, representing five ENL varieties (ICE Canada, Great Britain, Ireland, New Zealand and Australia), four ESL varieties (ICE Hong Kong, India, Philippines and Singapore) and one country where standard English is used as a second dialect alongside an English-based creole (ICE Jamaica). In terms of regional distribution, this means that the focus of the investigation will be on North American Englishes, Britain, Ireland, Southern Hemisphere settler varieties, on the one hand, and Asian Englishes on the other hand. One condition for selecting the ICE components to be included was that a complete sample (i.e. both the written and spoken part) had to be available, as one of the research questions for this paper was to include the factor 'medium/register' in the analysis, even though this meant that US English could not be included. At approximately 1 million words each, however, the ICE corpora are too small to

yield sufficient evidence on variability with individual triggers.⁷ More precisely, the overall number of variable contexts for the triggers turned out to vary greatly (see section 4.1). In a follow-up study, additional data were therefore retrieved from GloWbE to probe into the effect that the trigger might have on the choice of verb form in mandative sentences. GloWbE provides subcorpora of varying size for twenty varieties of English, sampling (informal) blog data, on the one hand, and ‘general’, on the other.⁸ In order to keep the amount of data for this follow-up study manageable, the focus in this part of the study is on AmE and BrE as reference varieties, as well as Indian English as an ESL variety.

3.2 Definition of the variable context

As pointed out above, previous studies have used different approaches in defining the variable context. This applies both to the selection of the triggering context and the variants that were included in the datasets. However, in order to allow for replicability, it is important to not only define the variable but also to provide sufficient detail on how data for the present paper were retrieved from the corpora. Johansson and Norheim (1988), for instance, used a predefined set of 17 suasive verbs, 11 nouns and 5 adjectives to retrieve mandative contexts from LOB and Brown, presumably on the basis of lists found in standard grammars. In the interest of replicability, this set of triggers was then used in diachronic studies of Brown-family corpora (e.g. Hundt 1998, Hundt & Gardner 2017, Waller 2017). Algeo (1992) lists 25 verbs, 29 nouns and 18 adjectives that may all trigger a mandative subjunctive, Crawford (2009) a total of 108 triggers and Kastronic and Poplack (2014) 240.⁹

⁷ Exact (and comparable) word counts for some ICE components are currently difficult to obtain. The online version of ICE components to be made available at URL will address this issue. For the current study, the dependent variable is defined as a choice context. This allows comparisons without normalization.

⁸ Note that the ‘general’ part of the sampling might occasionally also include blog data. For background on the corpus and the advantages and disadvantages of this resource for World Englishes research, see Davies and Fuchs (2016) and the responses to their article by Christian Mair, Joybrato Mukherjee, Gerald Nelson and Pam Peters in *English World Wide* 36(1).

⁹ Crawford’s list of triggers is based on an unpublished paper by Albakry and Crawford 2004, but the list provided in the appendix of Crawford (2009: 274) shows that it is not derived in a bottom-up fashion from corpus data but based on the lists provided in Quirk et al. (1985) and Övergaard (1995). Moreover, Crawford (2009) only uses the 16 nouns, 11 verbs and 6 adjectives that are attested at least once with the subjunctive in his news corpus (2009: 261). Kastronic and Poplack (2014) used previous studies as well as corpus data (following Hoffmann’s 1997 approach), but they fail to list the specific triggers that their study is based on. In a footnote they mention that *say, think, feel* initially were among the triggers but go on to say that they were not included in the rate calculations as they were overwhelmingly used with the indicative (Kastronic & Poplack, 2014: 77).

Hoffmann (1997) uses a corpus-based approach to arrive at the set of triggers he investigates: he first searched for instances of subjunctive *be*, and from the list of 276 items then used the most frequently attested triggers to retrieve his mandative sentences. However, even such an approach would miss instances like the following, that were retrieved searching for the sequence *that + personal pronoun + not*:

(4) On reviewing those remarks I *regret that I not provide* sufficient airing of Lessing's notions on that subject, and remained content with a cryptic reference. (ICE-SING, W2D-019)

(5) ... we pray and *pray* intensely *that we not be found wanting*. (ICE-PHI, S2B-023)

Ultimately, a complete set of mandative sentences is impossible to obtain from a corpus as mandative subjunctives can also be used without overt triggers (see Övergaard, 1995: 82). Moreover, in order to cover a relatively broad range of Englishes, the number of lexical triggers used in this paper had to be limited to keep the amount of data that had to be manually post-edited manageable. I opted to restrict data retrieval to eleven verbs and three adjectives.¹⁰ These were selected on the basis of having been shown in previous studies to regularly trigger subjunctives. For data retrieval from GloWbE, one hundred variable contexts were retrieved from the three sub-sections for the following verbs (all forms): *demand, order, propose, recommend, request* and *require*, giving a total of 1800 hits. Note that the deontic force of these triggers is quite different (i.e. stronger for *demand, order* and *require* than for *propose* and *recommend*). The former might therefore more frequently trigger a subjunctive than the latter.

Not all subordinate clauses following these triggers are instances of choice contexts, however. Waller (2017: 81) distinguishes four identifying contexts for the mandative subjunctive and introduces his own labels for them: a third person singular noun phrase followed by an unmarked verb (iNO-S), as in example (6), unmarked use of *be* (iBE) as in (7), an unmarked verb following a past-tense trigger (iST), as illustrated in example (8), and pre-verbal *not*-negation (iNEG), exemplified in example (9):¹¹

(6) May I *ask* that *the uh prosecutor furnish* us at least a copy of what they have furnished. (ICE-PHI, S1B-062)

¹⁰ The verbs (all variant forms) included were *ask, demand, dictate, insist, order, propose, recommend, request, require, suggest, urge*; the adjectives were *essential, imperative, important*.

¹¹ Kastronic and Poplack (2014: 73) simply define the subjunctive as a morphological category, i.e. as all instances of *be* and bare forms following a third-person singular subject NP. This means that their study did not include instances with disambiguating tense sequences as in examples (8) and (7).

- (7) Dewey *requires* that pupils *be* given wide opportunities for purposive inquiry. (ICE-IRE, S2B-035)
- (8) In short he *ordered* that *his troops use* terrorism. (ICE-CAN, W1A-005)
- (9) It's *important* that *you not wait* until severe soiling has occurred before cleaning your upholstery. (GloWbE, CA G)

These identifying contexts can also co-occur, as in the following example with a past tense in the main clause, pre-verbal negation and use of unmarked *be*:

- (10) ... he suggests opening up the least vulnerable part and *recommended* that all *visitors not be* concentrated on the same spot. (ICE-IRE, W1A-018)

However, in the following example with a past-tense trigger, the speaker self-corrects from a subjunctive to a past tense verb; I therefore decided not to include this particular instance:

- (11) And he *suggested that* as soon as I got here I came and *see saw* you. (ICE-GB, S1A-051)

Extensive manual post-editing of the concordances is necessary to include only true subjunctives and those instances with *should* that would also result in an unambiguous mandative had the modal been omitted. Thus, example (12) is ambiguous because the verb following the first-person subject is not *be* and could therefore be either indicative or subjunctive. Example (13) is ambiguous because the subject is a collective noun that is ambiguous in terms of number (collectives in English can be used with both plural and singular verbs and pronouns). Example (14) was excluded because there is no disambiguating context that would result in a mandative subjunctive were the modal to be left out.

- (12) ... what do you recommend *I read* to go into that more. (ICE-CAN, S1B-012)
- (13) The PQ is once again demanding that *Quebec opt out* of National Agricultural Program. (ICE-CAN, S1B-028)
- (14) *We propose* that *we should have* a meeting together to discuss the proposal further. (ICE-HK, W1B-023)

With collective nouns, pronominal clues in the immediate context can disambiguate the verb, making *sit up*, *use* and *allow* subjunctive; such instances were included in the final analysis:

- (15) She *demands* that *her audience sit up*, *use its* intellect and *allow* her plays to challenge. (ICE-NZ, W1A-002)

In addition to ambiguous instances, occurrences in which the trigger did not have mandative meaning were also removed during manual post-editing, including the following examples:

- (16) However the unions are *insisting* that the order *means* a return to the status quo of seven-member teams (ICE-JAM, S2B-005)
- (17) We have studies by psychologists about how jurors make up their mind *suggesting* [=‘implying’] *that* jurors typically decide by constructing or testing the alternative ... (ICE-GB, S2A-044)

I also decided to remove instances where the trigger was separated from the following context by a punctuation mark; typically, this concerns lists, as in the following example:

- (18) In his experiments with the culture of *tikog*, Herminio Pava, professor VI of Central Mindanao University in Musuan, Bukidnon, recommends the following:
*Planting distance should be at least 8 to 10 away from each other. (ICE-PHI, W2D-016-017)

ESL varieties provide additional cases that need to be carefully considered during manual post-editing. At times, the triggers are used with different semantics than we would expect from ENL varieties, as in the following example from ICE-IND, where *request* is used in the sense of ‘ask’.

- (19) So second September immedietly [sic!] I requested *them you send the <, > some of sample* which you've collected. (ICE-IND, S1B-029)

What complicates the analysis in this particular case is that the trigger is followed by what could either be interpreted as a subordinate clause (in which case *request* would be used as a transitive verb) or – since it is from the spoken part of the corpus – a quotation of what the speaker originally uttered as a request. For this reason, this particular occurrence was considered to be ambiguous and therefore not included in the analysis.

As pointed out in the introduction, English allows for the choice between a subjunctive, a modal construction or an indicative after mandative triggers. Other modals may also occur in mandative sentences, as the following examples show, including semi-modal *ought to*:

- (20) Mr Heseltine has been no less resolute than Mrs Thatcher in insisting that the dictator *must* withdraw from Kuwait. (ICE-GB, W2E-004)
- (21) We are simply asking that there *ought to* be a mechanism whereby our salary could be determined and adjusted in future ... (ICE-HK, S1B-042)

For the modelling of variation across the varieties (section 4.2) the envelope of variation was narrowed to include only the subjunctive and the modal construction with *should*. The

range of possible patterns that can follow a mandative trigger may, occasionally, even include a past tense indicative, as in (22) or (23).¹²

(22) and on the basis that there is apparently no new money available / *i would be loath to recommend* at this point / particularly as i'm departing / *that money was taken* from other schemes (ICE-NZ, S1B-07)

(23) After the doctor's report, the company asked that he *resigned* or *be fired*. (ICE-JAM, W2C-018)

The use of a past indicative in (22) may have to do with the fact that the trigger occurs in a hypothetical sentence and the subordinate clause occurs at quite a distance from the trigger.

3.3 Predictor variables

In a second step, the mandative sentences retrieved from the ICE corpora were coded for two external predictor variables (variety and medium/register) and six contextual variables. The factor 'variety' is self-explanatory, but the external predictor 'medium/register' needs further commentary. The ICE sampling frame distinguishes a total of eight written and four spoken categories (see e.g. Greenbaum 1996). For the medium/register analysis in this paper, however, the four macro-categories 'spoken dialogue', 'spoken monologue', 'written unpublished' and 'written published' were used to avoid the risk of having too sparsely populated cells in the subsequent statistical modelling. The contextual linguistic variables included in the data coding were 'trigger' (with the values of the individual lexical items), 'trigger type' (verb vs. adjective), 'controlling subject' (with the values 'third person singular' vs. 'non-third person'), 'verb' (for the lexical verbs in the subjunctive/modal construction, with the values 'be' vs. 'other'), 'negation' (with the values 'negative' vs. 'affirmative') and 'subordination' (with the values 'that' vs. 'zero').

3.4 Data retrieval

3.4.1 Core study

In order to include instances where the subordinating conjunction had been omitted (see examples (24)-(26)), data retrieval relied on the mandative triggers, only.

(24) Initially the captors demanded \emptyset the cash *be dropped* by plane (ICE-AUS, S2B-005)

(25) I insisted \emptyset he come and see our establishment before he did that. (ICE-IND, W2F-009)

¹² The past indicative is followed by a *be*-subjunctive in example (23); the latter was included in the analysis.

(26) ... the council's planning consultant recommended \emptyset the consent *exclude* general engineering fibreglassing spray painting and steam or sand blasting operations. (ICE-NZ, W2C-008)

One disadvantage of this approach is that it retrieves a relatively large number of false positives. For the lemma *ASK*, for instance, the search yields well over 2,000 hits, of which only 37 showed variation between *should* and a subjunctive. Numerous instances were either not mandative uses of *ASK* (see (27)) or avoided the choice between the subjunctive and a modal construction by using a non-finite complement (as in (28) and (29)).¹³

(27) I asked Malang *what he thinks* the peoples of the planet should do to strive for sustainability. (ICE-CAN, S2B-038)

(28) We would certainly ask *for* those *to reduce* the time. (ICE-GB, S1A-024)

(29) Then after the interview they asked me *to join* in the month of July. (ICE-IND, S1A-035)

Complementation of the trigger *URGE* provides further interesting evidence of the possibility of avoiding the choice between a subjunctive and modal periphrasis. In addition to combining an agent NP with a non-finite construction as in (30), it is also possible to use a deverbal noun and move the agent to a *by*-phrase, as in (31).

(30) I would urge *members to give consideration* to nominating as a Director. (ICE-AUS, W1B-024)

(31) He said problems with boarding houses were highlighted by the Burdekin report on the rights of mentally ill people, which urged *stringent licensing and regulation by State governments*. (ICE-AUS, W2C-017)

Statistical analysis of the ICE data showed that the trigger is a significant factor (see section 4.2), but one that needed to be further controlled as the raw number of variable contexts per trigger in ICE varied greatly.

3.4.2 Follow-up study

For the follow-up study, the goal was to retrieve 100 variable contexts per trigger and variety. This meant that the number of triggers and varieties had to be reduced to keep the amount of data manageable. The focus will be on AmE and BrE as metropolitan reference varieties and IndE as an institutionalised second-language variety. The six triggers (*demand*, *order*, *propose*, *recommend*, *request* and *require*) were selected because they were

¹³ Note that in some ESL varieties and JamE, *request* is also regularly used in the sense of 'ask' with the same complementation pattern, i.e. a direct object followed by a *to*-infinitive rather than a *that*-complement clause: "I request you to convey your sanction at the earliest" (ICE-IND, W1B-021).

relatively frequently attested in the variable context in previous research.¹⁴ In order to increase precision and reduce the number of sentences that had to be removed at the manual post-editing stage, data extraction for the follow-up study was limited to instances where the trigger was followed directly by a subordinating conjunction. The GloWbE data were manually post-edited and coded for five predictor variables: ‘variety’, ‘trigger’, ‘controlling subject’, ‘verb’ and ‘negation’. A predictor variable ‘register’ (general vs. blog) was not included as about 20% of the material included in the ‘general’ category also comes from blogs (see Davies & Fuchs, 2015: 4).

4. Results: Contextual and/or linguistic factors?

I will first present an overview of the distribution of subjunctives and periphrastic constructions with *should* in the ICE data, then move on to the statistical analysis of this data set (sections 4.1 and 4.2). The additional evidence from GloWbE on the potential impact that the lexical triggers may have on the choice of a subjunctive will be presented in section 4.3.

4.1 Distribution of mandative subjunctives in ICE by predictor variable

The search for mandative sentences with an unambiguous mandative verb or periphrastic *should* construction in the 10 ICE corpora yielded a total of 403 instances, of which the majority (70.7%) are subjunctives. Not all regional varieties contribute to the dominance of the subjunctive, though, as the distribution of the variants across the WEs examined (see Table 1) shows: While CanE shares the preference with its American neighbour for the subjunctive, periphrastic *should* is preferred in ICE-GB. In the ICE data, NZE and AusE also clearly prefer the subjunctive. As in previous studies (Schneider 2005, 2011; Collins et al. 2014), PhilE aligns with the ‘American’ usage pattern in its strong preference for the subjunctive. Varieties showing a more even distribution of the two options are IrE, IndE and HKE.

	CanE	BrE	IrE	NZE	AusE	HKE	IndE	PhilE	SingE	JamE
<i>should</i>	6	16	18	15	5	16	16	11	10	5
subjunctive	35	10	18	47	35	14	16	46	26	38

¹⁴ Data were originally also collected for *urge*, but since these did not amount to the required number of 100 variable contexts per variety, these were not included.

Table 1: Distribution of subjunctives and should-constructions across varieties (raw frequencies) in ICE

With respect to the distribution of subjunctives against modal variants across speech and writing (see Table 2) the ICE data do not reveal a marked difference. Interestingly, the informal spoken and written registers yield an even higher proportion of subjunctives than the formal ones, but this is not a finding that finds a straightforward explanation, e.g. in the innovative varieties providing proportionally more evidence from informal spoken and written contexts. The fact that the largest total number of choice contexts comes from written published English probably explains why people associate the subjunctive with this register.

	dialogue	monologue	unpublished	published
<i>should</i>	16 (22.5%)	41 (37.6%)	13 (20.6%)	48 (30%)
subjunctive	55 (77.5%)	68 (62.4%)	50 (79.4%)	112 (70%)

Table 2: Distribution of subjunctives and *should*-constructions across ICE registers (macro-categories)

Preliminary evidence from ICE indicates that the lexical trigger plays an important role in the choice between the two variants (see Figures 2a and 2b), with *urge*, *request*, *dictate* or *imperative* favouring the subjunctive and *recommend* or *suggest* showing a more even distribution of the variants. However, as the amount of data for individual triggers varies greatly, additional evidence on this factor for a limited number of varieties is discussed in section 4.3 below.

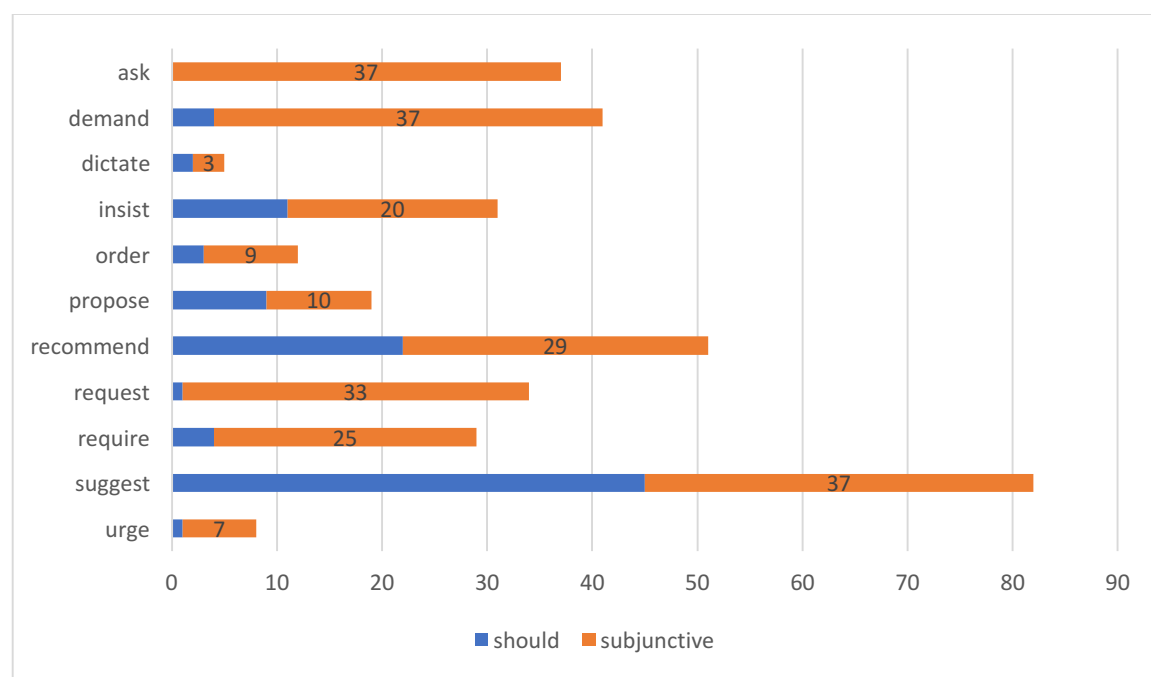


Figure 2a: Distribution of *should* vs. subjunctive (raw frequencies) across verbal triggers in ICE

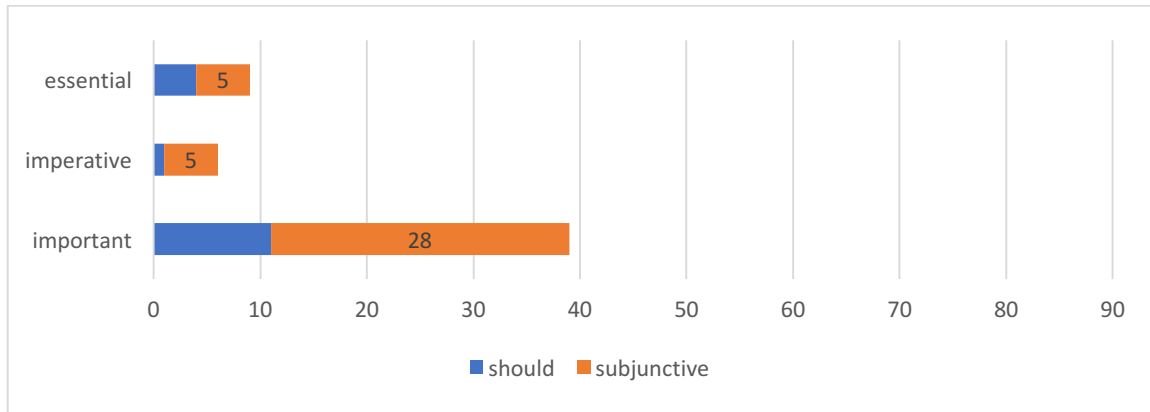


Figure 2b: Distribution of *should* vs. subjunctive (raw frequencies) across adjectival triggers in ICE

When we look at the role that the controlling subject may play (see Table 3), we see that the third person is an obvious stronghold of the subjunctive (61% of all subjunctives occur with a third person subject). However, in terms of choice contexts, non-third person subjects proportionally occur more often with the subjunctive than with a modal.

	third person singular	other
<i>should</i>	80 (31.5%)	38 (25.5%)
subjunctive	174 (68.5%)	111 (74.5%)

Table 3: Distribution of subjunctives and *should*-constructions by controlling subject in ICE

With respect to the verb used in the choice context, the summary statistics in Table 4 show that, in absolute terms, *be* is a stronghold for the subjunctive, as in previous research; however, proportionally, the choice for the subjunctive is not more likely with *be* than with other verbs.

	<i>be</i>	other
<i>should</i>	90 (32%)	28 (23%)
subjunctive	191 (68%)	94 (77%)

Table 4: Distribution of subjunctives and *should*-constructions by variable verb (*be* vs. other)

As far as the possibility of *that*-omission is concerned, Table 5 shows that overall it is relatively infrequent at 11.2% of all mandative contexts in the ICE data. More importantly, there is no substantial difference between subjunctives and periphrastic constructions with *should* as far as this factor is concerned. Interestingly, some regional varieties allow for more *that*-omission (New Zealand, Australian and Philippine English) whereas others strongly favour overt subordination in mandative constructions (namely British, Irish and Jamaican English).

	<i>should</i>	subj.	CanE	BrE	IrE	NZE	AusE	HKE	IndE	PhilE	SingE	JamE
<i>that</i>	107 (90.7)	251 (88.1)	38 (32)	25 (10)	36 (18)	49 (37)	31 (26)	25 (11)	30 (14)	49 (41)	33 (24)	42 (38)
zero	11 (9.3)	34 (11.9)	3 (3)	1 (0)	0 (0)	13 (10)	9 (9)	5 (3)	2 (2)	8 (5)	3 (2)	1 (0)

Table 5: Distribution of subordination (zero vs. *that*) across varieties (figures in brackets for varieties give the subjunctives, only)

Finally, the ICE data show that the subjunctive strongly disprefers negative contexts: out of 285 subjunctives, only 3 (1.1%) were negated, compared with 12 (10.2%) negated instances of *should* (out of a total of 118) in mandative constructions. In addition to the one attestation of a negated subjunctive ICE-IRE quoted above (see example (10)) there is one each in ICE-CAN and ICE-JAM:

- (32) Since this story is popular with men, it is *important that* the woman *not become* too independent. (ICE-CAN, W1A-017)
- (33) Oh, but the crab catchers have *suggested that you not try* that method, have they? (ICE-JAM, W2F-006)

Corpus data from ICE thus confirm previous corpus-based finding (e.g. Hundt 1998, Leech et al. 2009), which also showed negated subjunctives to be rare. Moreover, all three instances are from the written part of the corpus. But both example (33) and the following instance from a chatty article in the Indian edition of *Cosmopolitan* magazine show that negated subjunctives are not necessarily limited to formal contexts in writing:

- (34) If you are one of those women who complain about not finding a guy who knows how to cook and clean, chances are you aren't hanging out with the right kind of man. In all honesty, tradition *demandd that* Indian men *not step* into the kitchen but if you are going to break every other kind of tradition then the least you can do is not pick this one as the only tradition to maintain. (GloWbE, cosmopolitan.in)

The question remains, however, what the relative importance of the factors is and how usage of subjunctives plays out across the different regional varieties.

4.2 Statistical analysis: variable importance

A traditional approach to modelling variable importance in a multivariate approach is through variable rule or regression analysis. There are two problems inherent in logistic regression models, however, as Szmrecsanyi et al. (2016) point out, namely predictor multicollinearity (the risk that predictors in the model are correlated) and data overfitting. In order to model probability grammar in World Englishes, they use both a random forest

analysis and a conditional inference tree (ctree) analysis, an approach also recommended by Tagliamonte and Baayen (2012). The random forest analysis is a variant of permutation testing which has the advantage that it does not assume a certain data distribution but instead builds the distribution by recursively resampling the data. It provides information (on the basis of a large number of trees) on the relative importance of predictor variables but does not indicate how these might interact with each other (e.g. Strobl, Malley & Tutz 2009 or Strobl, Hothorn & Zeileis 2009). The conditional inference tree (ctree) analysis returns a single tree which, on the basis of recursive partitioning of the underlying data, makes predictions in the form of binary splits of the data in a hierarchical fashion, thus showing predictor interaction. For the purposes of this study, interaction between ‘variety’ and other predictor variables is of particular interest. The two different kinds of analysis thus complement each other. They were computed using R’s party and partykit packages.

The relative ranking of the factors in the random forest analysis (with $n_{tree} = 500$ and $m_{try} = 2$) in Figure 3 returns ‘trigger’ as the most important predictor variable in the choice between subjunctive and *should* periphrasis, with ‘variety’ coming second. All the other factors are far less important.

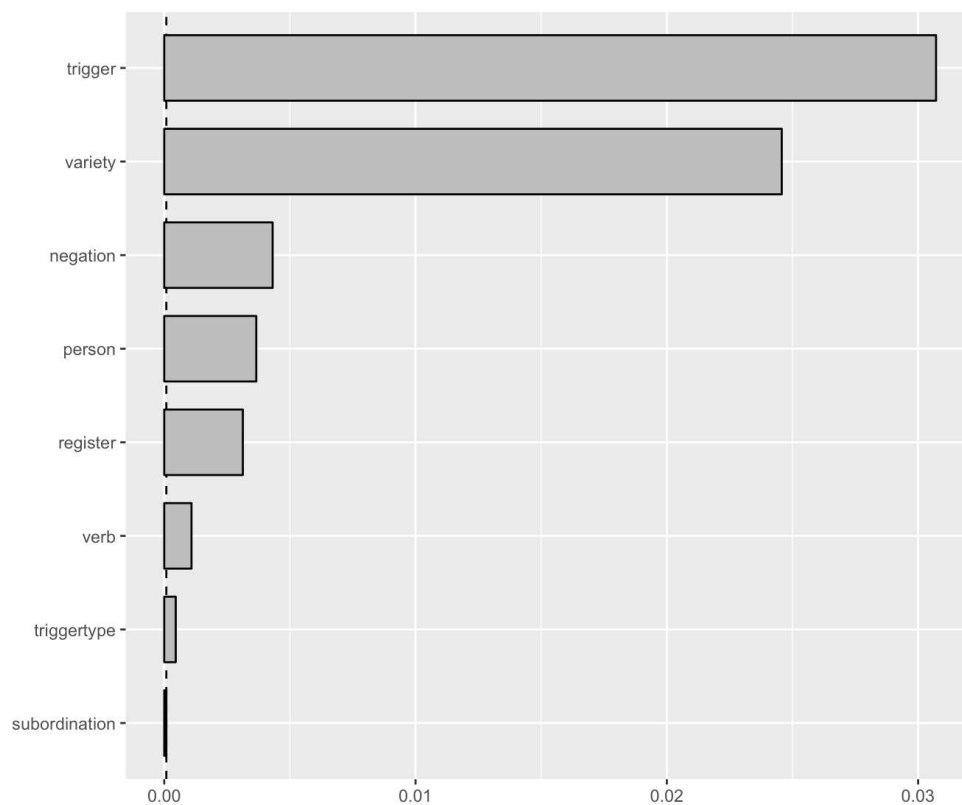


Figure 3: Variable factor importance (Random Forest Analysis) for the ICE data

The variable importance plot in Figure 3 is based on the run with all predictor variables. Testing for model fit with Somers2 Dxy returns a prediction accuracy of 0.75 and a C-index value of 0.875, which is above the 0.8 level recommended e.g. in Tagliamonte & Baayen (2012: 156) thus already indicating a good model fit. The best model fit (with a C-index of 0.896 and a prediction accuracy of 0.79) was returned in a run with only ‘trigger’, ‘variety’ and ‘register’ and ‘person’ as predictor variables.

The single tree (with maxdepth = 4, mincriterion = 0.95) also selects ‘trigger’ at the first split, but only ‘variety’ as the other predictor variable (note, however, that prediction accuracy for the ctree is at 0.425 and thus slightly below the 0.5 level, with a C-index of 0.71). Figure 4 also shows that the triggers *ask, demand, imperative, order, request, require* and *urge* strongly favour the subjunctive (regardless of variety), whereas triggers *dictate, essential, important, insist, propose, recommend* and *suggest* are predominantly followed by a subjunctive in AusE, CanE, JamE, NZE, PhilE and SingE, whereas BrE, HKE, IndE and IrE show a preference for the modal construction with this set of triggers.

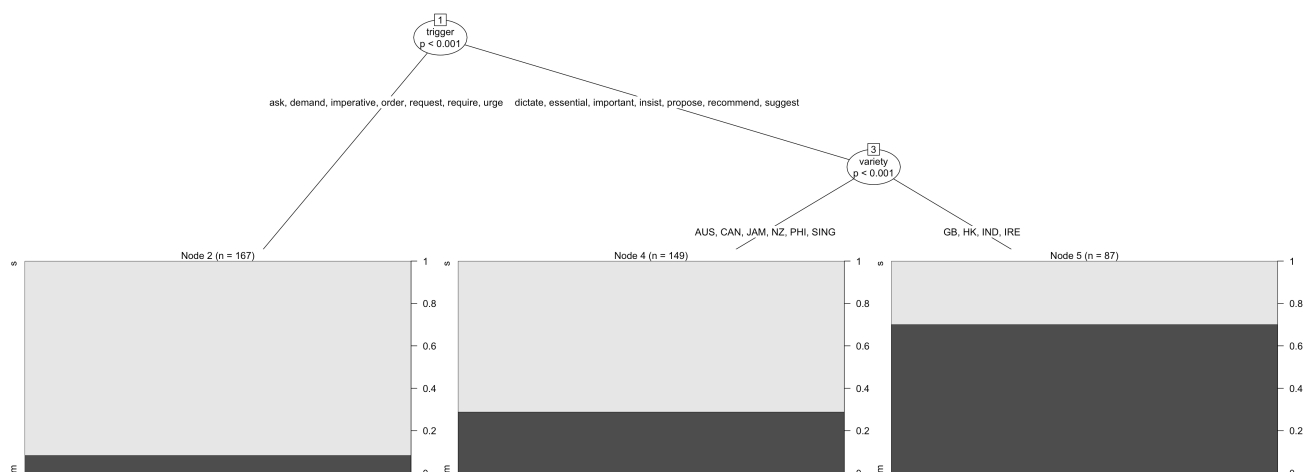


Figure 4: Plotting predictor interaction for mandates in ICE (ctree)

4.3 Gauging the effect of the trigger

Figure 5 summarizes the distribution of subjunctives and *should*-constructions across the three varieties in the GloWbE data. As expected, AmE shows the highest proportion of subjunctives. Interestingly, the proportion of subjunctives in the GloWbE is higher for BrE and IndE than previous studies, based on standard reference corpora, would suggest. This could be the result of the continued spread of the subjunctive in these varieties, as the GloWbE data are more recent than the material sampled in the Brown family and ICE corpora. Alternatively, it might have to be attributed to the nature of the data included in

GloWbE (i.e. the sampling frame, while trying to emulate the ICE criteria, might not have succeeded in providing a stylistically comparable set of data). Overall, the differences between AmE and BrE still prove significant at $p < 0.01$ in a chi-square test, whereas those between BrE and IndE do not.

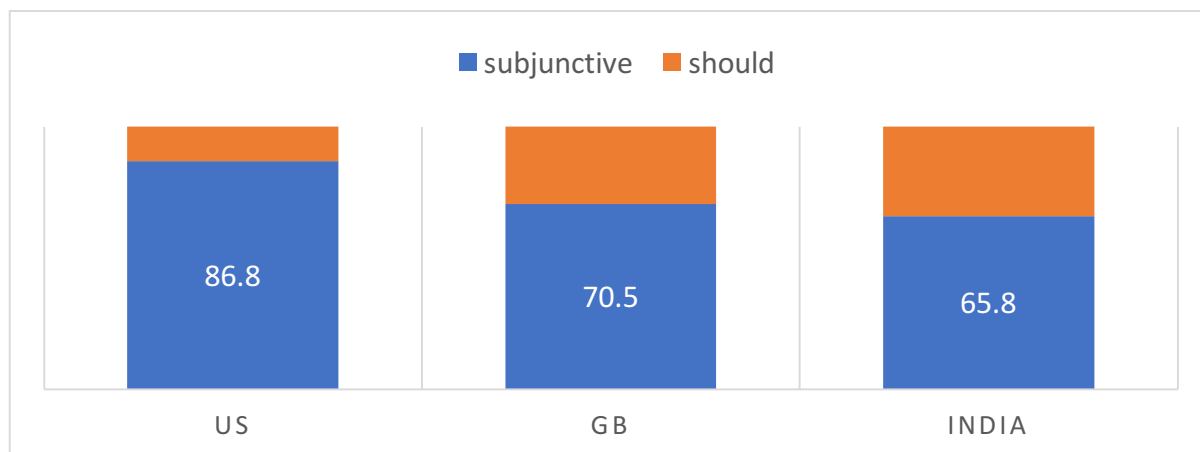


Figure 5: Mandative subjunctives (%) in GloWbE (600 variable contexts each per variety)

If we turn to the distribution by lexical trigger (see Figure 6), we see that there are verbs that strongly prefer the subjunctive across the GloWbE data analysed here, namely *require* (88.7%) and *request* (91.6%), to a somewhat lesser extent *demand* (83%); *recommend* and *order* trigger subjunctives at lower rates, namely at 67.3% and 68.3%, respectively; *propose*, finally, has the highest number of periphrastic *should*-constructions at just over 50% (52.6%). The lexical trigger thus turns out to have a significant effect on the choice of a subjunctive over a periphrastic construction in a larger dataset that provides samples of equal size per trigger.

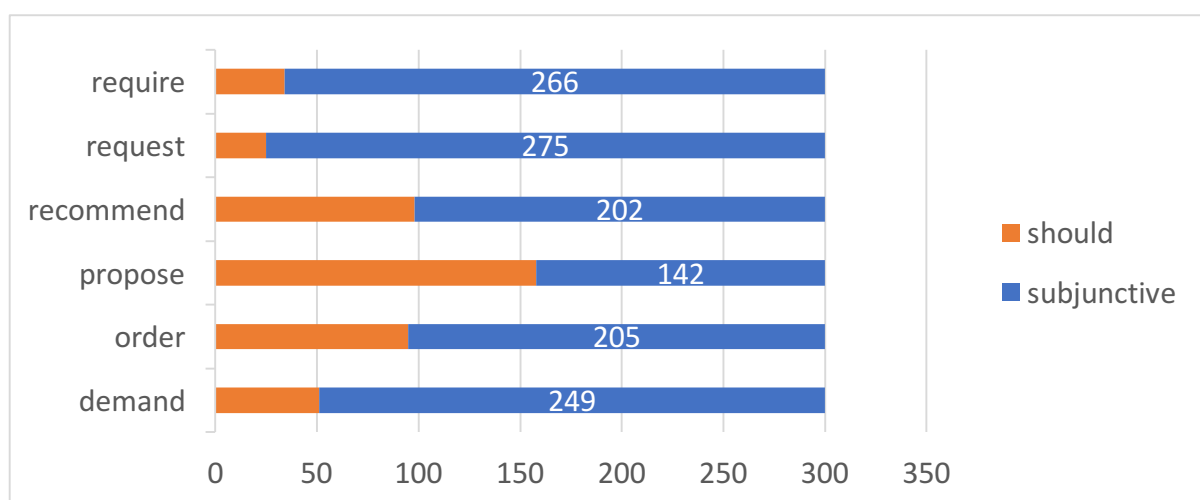


Figure 6: Mandative subjunctives (%) in GloWbE (300 variable contexts per verb, 100 from each variety)

As in the ICE data, the verb *be* is not used significantly more often as a subjunctive than other verbs (the proportion of subjunctives for *be* and other verbs is 74%); negation does not prove significant in a chi square test, either (overall, negative subjunctives only amount to just under 4% of all subjunctives). Third person subjects, finally, are followed significantly more often by a subjunctive (76.7%) than other subjects (70.4%) in the GloWbE data.

The random forest analysis (with *ntree* = 500 and *mtry* = 2) shows that ‘trigger’ is the most important predictor variable, with ‘variety’ coming second (see Figure 7). Model validation with *somers2* returns an accuracy level of 0.546 and a C-index of 0.773 (i.e. above the 0.5 level of random assignment but very slightly below the 0.8 level recommended by Tagliamonte & Baayen, 2012), i.e. shows that the analysis with all predictor variables already results in a good model fit. The best model fit is achieved with ‘trigger’, ‘variety’, ‘verb’ and ‘person’ as predictor variables (accuracy level of 0.558 and a marginally higher C-index of 0.779).

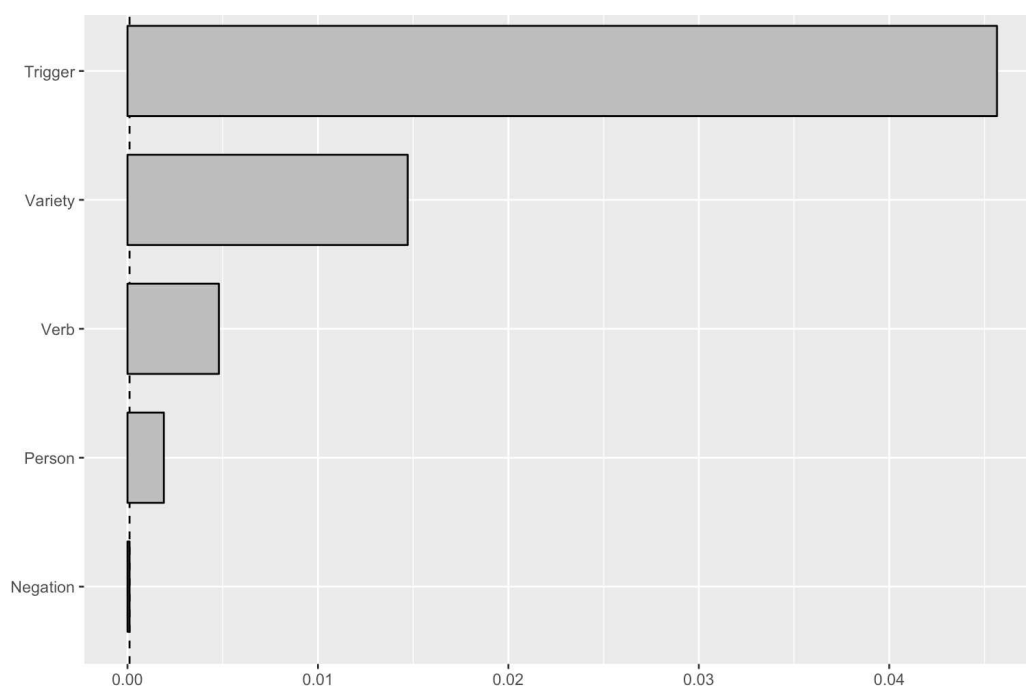


Figure 7: Variable factor importance (Random Forest Analysis) for the GloWbE data

The single tree in Figure 8 based on the GloWbE data and including all predictor variables splits the data into two groups: triggers that strongly favour the subjunctive (nodes 2-4) vs. those that do so to a lesser extent, particularly so in BrE and IndE (nodes 13-15).

Interestingly, AmE and BrE show a similarly strong preference for subjunctives with the triggers *demand*, *request* and *require* (nodes three and four), whereas with verbs that trigger the periphrastic construction more often (i.e. *order*, *recommend* and *propose*), BrE

and IndE align. In other words, IndE is the most conservative variety throughout, even with verbs that strongly lean towards the use of the subjunctive (see node 2), AmE is the most advanced, whereas BrE takes up an intermediate position. The late selection of the factors ‘person’ and ‘verb’ (nodes 14 and 15) indicates that these factors are less important than ‘trigger’ and ‘variety’, and fits in well with the random forest analysis. Finally, ‘negation’ is not selected in the single tree, either.

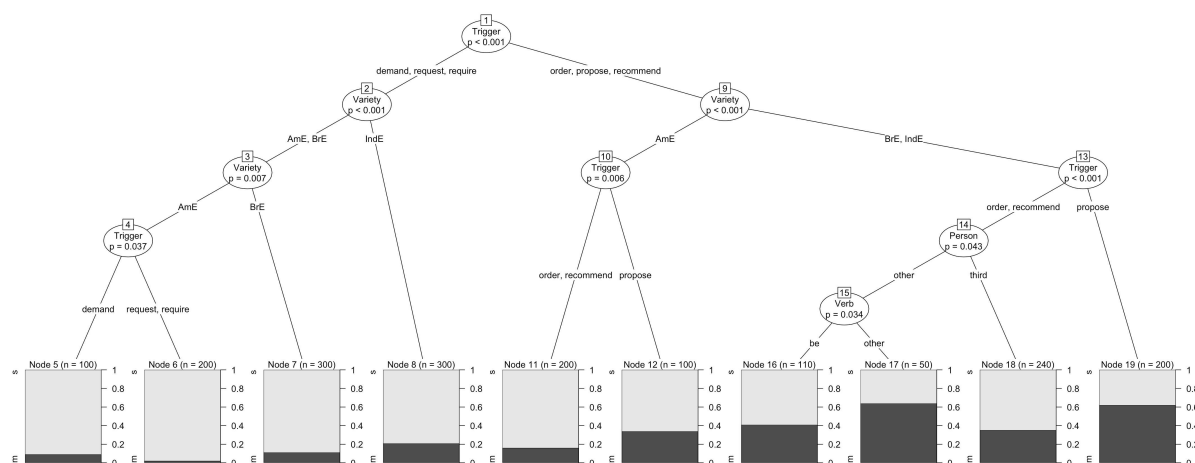


Figure 8: Plotting variable importance (trigger and variety) in GloWbE data

5. Discussion: Mandative subjunctives and models of World Englishes

An important finding of the present study is that, on the basis of the ICE data, there is no significant effect of register on the choice between a subjunctive and the periphrastic construction with *should*. This fits in with previous research (Szmrecsanyi et al. 2016: 133), which also found that ‘variety’ consistently ranked higher than ‘register’ as a variable predicting the choice for the dative and genitive alternations as well as particle placement across World Englishes. The fact that variability is strongly determined by the lexical trigger also fits in well with previous research in that lexico-grammatical variation typically shows up as an indicator of indigenization (see e.g. Schneider 2007), in this case ‘probabilistic indigenization’ in the sense of Szmrecsanyi et al. (2016):

... the process whereby stochastic patterns of internal linguistic variation are reshaped by shifting usage frequencies in speakers of post-colonial varieties. To the extent that patterns of variation in a new variety A, e.g. the probability of item x in context y, can be shown to differ from those of the mother variety, we can say that the new pattern represents a novel, if gradient, development in the grammar of A. These patterns need

not be consistent or stable ..., but they nonetheless reflect the emergence of a unique, region-specific grammar.

The results from the ICE corpora reveal that variation in mandative sentences cuts across ENL, ESL and ESD varieties¹⁵: ESL IndE aligns closer to BrE than to SingE, another ESL variety, for instance. While the tripartite distinction does not help explain the variability in this area of morphosyntax, neither does a genetic model (e.g. Strevens 1992), that groups varieties into how they derive from BrE and AmE as matrilects. Variability in the use of mandative subjunctives also cuts across these 'genetic' distinctions: neither NZE nor SingE align with their 'matrilect' (BrE).

Theoretically, language contact might provide an explanation as to why the subjunctive is preferred in SingE: since the mandative subjunctive is identical in form with the base form of the verb, speakers with a variety of Chinese as their first language might prefer unmarked verbs seeing that Chinese does not mark the subjunctive on the verb. The subjunctive, in this interpretation, would also fit in well with a general tendency in contact varieties of English towards simplification. However, this seemingly straight-forward explanation starts to crumble when we look at HKE, which aligns with BrE in its preference for a periphrastic *should* construction in mandative sentences, despite the fact that it shares its main substrate (varieties of Chinese) with SingE.

Traditional models of WEs have been criticized for being rather static. So an obvious question is whether a more dynamic model, such as Schneider's (2003, 2007) is better suited to account for the variability found in mandative subjunctive use. Figure 9 lists the postcolonial varieties according to the developmental stage they occupy.¹⁶

¹⁵ The distinction, according to Schneider (2017: 39), was introduced by Quirk et al. (1972).

¹⁶ Irish English is not typically discussed as an example of a postcolonial English within Schneider's Dynamic Model, and does not seem to fit it easily, among other things because of its much longer history of colonization (i.e. since the middle ages). Ronan (2016) argues that IrE can be said to have reached Phase 5 (differentiation), following a somewhat different trajectory of change. While SingE may have progressed further towards stage 5 and PhilE towards stage 4 since the publication of Schneider's (2007) monograph, data collection for the ICE corpora took place in the last decades of the twentieth century, making the original classification in Schneider (2007) a suitable point of comparison.

	Phase 2	Phase 3	Phase 4	Phase 5
CanE				-----
NZE				-----
AusE				-----
IrE				-----
SingE			-----	
HKE		-----		
PhilE		-----		
IndE		-----		
JamE			-----	

Figure 9: Developmental phase of post-colonial varieties of English according to Schneider's (2007) model (CanE = Canadian English; NZE = New Zealand English; AusE = Australian English; IrE = Irish English; SingE = Singapore English; HKE = Hong Kong English; PhilE = Philippine English; IndE = Indian English; JamE = Jamaican English; classification on the basis of Schneider 2007 and Ronan 2016)

One of the basic tenets of Schneider's model is that, as a variety progresses along the developmental cycle, it becomes nativized and thus diverges, structurally, from its matrilect. For lexicogrammatical features (e.g. collostructional variation), greater distance from the input variety has been shown to tally well with developmental stage (see e.g. Mukherjee & Gries 2009), albeit with register as a factor that is likely to add 'noise' to recorded variation in that nativization is more palpable in spoken than in written registers (see Gries & Mukherjee 2010: 542 or van Rooy 2010).¹⁷ The model would account for the relatively conservative nature of HKE, PhilE and IndE (as adhering to their respective matrilects during stage 3)¹⁸ as well as the divergence of SingE away from the BrE model (during stage 4). However, we would also expect to see divergent development for IrE (phase 5) and JamE (phase 4), but both varieties closely align with their respective matrilects.

AmE, the result of what Mesthrie (2006: 388) calls 'the third crossing', has moved beyond phase 5 in Schneider's Dynamic Model and risen to a global standard competing with the original matrilect from the British Isles: "globalization seems to be propelling US English into a position as a potential rival to standard southern British English" (ibid.). An

¹⁷ We saw above that the spoken vs. written mode do not play a significant role in this case.

¹⁸ Sedlatschek (2009: 287) reports that local textbooks and usage guides explicitly endorse the periphrastic construction and warn against the modal variant as being too formal.

alternative approach might therefore be to simply view the observed variation as a result of Americanisation, an interpretation which would also fit in with Mair's (2013) model of World Englishes in which AmE serves as a hub variety for ongoing change across Englishes.¹⁹ Americanisation has been proposed as an explanation in previous research on the mandative subjunctive, e.g. to explain the increase in BrE (Kjellmer 2009: 256). It has also been claimed to play a role, more generally, in ongoing change in CanE (e.g. Boberg 2004), AusE (e.g. Collins 2009) and NZE (e.g. Bayard 1999, Green & Bayard 2000). While for SingE, Americanisation has been observed in the areas of phonology (Tan 2015) or spelling and vocabulary (Hänsel & Deuber 2013), the case seems to be less clear for a variety like IndE (see e.g. Cowie 2007 on the development of a regionally 'neutral' accent in call centres).

One problem with the concept of Americanisation is that it is difficult to verify, at least if defined as a process that can be traced to speakers' conscious choices. Vine (1999), for instance, shows that the use of 'American' lexical items in New Zealand does not necessarily go hand in hand with speakers' awareness of them as being Americanisms. Similarly, Hundt (1998: 94) reports on evidence from a small-scale elicitation experiment that showed how informants in New Zealand were unaware of the subjunctives' association with AmE. This confirmed a previous claim by Algeo (1992: 603-604), who suspected that speakers were unlikely to be aware of the subjunctive's regional association with AmE and the periphrastic construction's connection with BrE:

Americans are aware that the British talk differently, but if asked to specify particular forms of difference, few could cite any more than a few hoary old chestnuts like British *lift* for *elevator*. That any except grammarians would have an awareness of the mandative subjunctive is highly improbable.²⁰

Thus, while 'Americanisation' might be an attractive explanatory concept for the distribution of variants in mandative sentences across the varieties in ICE, it is unlikely that speakers in Australia, New Zealand and Singapore are either aware of this as being an

¹⁹ This is the more wide-spread use of the term. Mufwene (2009: 365) uses it to refer to the indigenization of English in North America: "I am using the term **Americanization** here in the sense of 'becoming American in character'. In the case of English, it means becoming different from British varieties by acquiring characteristics that make it particularly American."

²⁰ Note that, while there is some discussion of the mandative's association with AmE on some language-related websites, such as the *English Language & Usage* forum or the advice page *Perfect your English*, these can be considered 'expert' discussions of the topic (see <https://english.stackexchange.com/questions/76578/why-is-american-english-so-wedded-to-the-subjunctive> and <http://www.perfectyourengish.com/writing/american-and-british-grammar.htm>).

Americanism nor that they are consciously adopting the mandative subjunctive in order to sound more American. The question is whether ‘Americanisation’ is therefore the right choice of term, in the first place.²¹ Moreover, Americanisation – if at work – might be working alongside other areal factors. Peters (2009: 135), for instance, accounts for the relatively high proportion of mandative subjunctives in both PhIE and SingE as being possibly the result of regional alliances (i.e. SEAMEO, the South-East Asian Ministers of Education Organization). If anything, this possibility would lend support to yet another model of World Englishes as a network of local centres which speakers might be relating to. However, empirically verifying norm orientation to a more local standard than the traditional metropolitan standards of BrE and AmE using corpus data as the only source of evidence is just as difficult as proving ‘Americanisation’ (see e.g. Hundt 2013).

References

Corpora

- ICE *The International Corpus of English*
 [URL to be provided once it’s been migrated to the new address]
- GloWbE Corpus of Global Web-based English (<http://corpus.byu.edu/glowbe/>)

Bibliography

- Algeo, John. 1992. British and American mandative constructions. In Claudia Blank, ed. *Language and Civilization: A Concerted Profusion of Essays and Studies in Honour of Otto Hietsch*, vol. 2, Frankfurt: Lang, 599-617.
- Bayard, Donn. 1999. Getting in a flap or turning off the tap in Dunedin? Stylistic variation in New Zealand English intervocalic (-T-). *English World-Wide* 20(1), 125-155.
- Boberg, Charles. 2004. English in Canada: phonology. In Kortmann, Bernd, Edgard W. Schneider, Kate Burridge, Rajend Mesthrie and Clive Upton, eds. *A Handbook of Varieties of English*. Volume 1: Phonology. Berlin: de Gruyter, 351-365.
- Cheshire, Jenny, Paul Kerswill, Sue Fox and Eivind Torgersen. 2011. Contact, the feature pool and the speech community: The emergence of Multicultural London English. *Journal of Sociolinguistics*, 15(2): 151–196.
- Collins, Peter. 2009. Modals and quasi-modals. In Pam Peters, Peter Collins and Adam Smith, eds. *Comparative Studies in Australian and New Zealand English: Grammar and Beyond*. Amsterdam: Benjamins, 73-87.
- Collins, Peter. 2015. Diachronic variation in the grammar of Australian English. Corpus-based explorations. In *ibid.*, ed. *Grammatical Change in English World-Wide*. Amsterdam: Benjamins, 15-42.

²¹ For a careful discussion of the methodology and theoretical underpinnings needed to ‘prove’ Americanisation/globalisation, see Meyerhoff and Niedzielski (2003).

- Collins, Peter, Ariane Macalinga Borlongan, Joo-Hyuk Lim and Xinyue Yao. 2014. The subjunctive mood in English. A diachronic analysis. In Simone E. Pfenninger, Olga Timofeeva, Anne-Christine Gardner, Alpo Honkapohja, Marianne Hundt and Daniel Schreier, eds. *Contact, Variation and Change in the History of English*. Amsterdam, Benjamins, 259-280.
- Cowie, Claire. 2007. The accents of outsourcing: the meanings of 'neutral' in the Indian call centre industry. *World Englishes* 26(3): 316-330.
- Crawford, William J. 2009. The mandative subjunctive. In Günter Rohdenburg and Julia Schlüter, eds. *One Language, Two Grammars? Differences between British and American English*. Cambridge: Cambridge University Press, 257-276.
- Davies, Mark and 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.
- Denison, David. 1998. Syntax. In Suzanne Romaine, ed. *The Cambridge History of the English Language, Volume IV, 1776-1997*. Cambridge: Cambridge University Press, 92-329.
- Fowler, Henry. 1926. *A Dictionary of Modern English Usage*. Oxford: Clarendon Press.
- Green, James A. and Donn Bayard. 2000. Zed to zee: notes from Dunedin on the Americanisation of the NZE lexicon through time. *Te Reo: Journal of the Linguistics Society of New Zealand* 43: 63-70.
- Greenbaum, Sidney, ed. 1996. *Comparing English Worldwide: The International Corpus of English*. Oxford: Clarendon.
- Gries, Stephan Thomas and Joybrato Mukherjee. 2010. Lexical gravity across varieties of English. An ICE-based study of n-grams in Asian Englishes. *International Journal of Corpus Linguistics* 15(4): 520-548.
- Hänsel, Eva Canan and Dagmar Deuber. 2013. Globalization, postcolonial Englishes, and the English language press in Kenya, Singapore, and Trinidad and Tobago. *World Englishes* 32(3): 338-357.
- Harsh, Wayne. 1968. *The Subjunctive in English*. Tuscaloosa AL.: University of Alabama Press.
- Hoffmann, Sebastian. 1997. *Mandative Sentences. A Study of Variation on the Basis of the British National Corpus*. Zürich: Unpublished Lizentiats-Thesis.
- Hundt, Marianne. 1998. It is important that this study (*should*) be based on the analysis of parallel corpora: on the use of the mandative subjunctive in four varieties of English. In Lindquist, Hans, Staffan Klintborg, Magnus Levin and Maria Estling, eds. *The Major Varieties of English*. Växjö: Växjö University, 159-75.
- Hundt, Marianne. 2013. The diversification of English: Old, new and emerging epicentres. In Daniel Schreier and Marianne Hundt, eds. *English as a Contact Language*. Cambridge: Cambridge University Press, 182-203.
- Hundt, Marianne and Anne Christine Gardner. 2017. Corpus-based approaches: Watching English change. In Laurel Brinton, ed. *English Historical Linguistics: Approaches and Perspectives*. Cambridge: Cambridge University Press, 96-130.
- Johansson, Stig and Else Helene Norheim. 1988. The subjunctive in British and American English. *ICAME Journal* 12: 27-36.
- Kastronic, Laura and Shana Poplack. 2014. The (north) American English mandative subjunctive in the 21st century: Revival or remnant? *University of Pennsylvania Working Papers in Linguistics* 20(2): 71-80.

- Kjellmer, Göran. 2009. The revived subjunctive. In Günter Rohdenburg and Julia Schlüter, eds. *One Language, Two Grammars? Differences between British and American English*. Cambridge: Cambridge University Press, 246-256.
- Leech, Geoffrey, Marianne Hundt, Christian Mair, and Nicholas Smith. 2009. *Change in Contemporary English: A Grammatical Study*. Cambridge: University Press.
- Mair, Christian. 2006. *Twentieth-Century English: History, Variation and Standardization*. 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.
- Mair, Christian. 2017. World Englishes and corpora. In Markku Filppula, Juhani Klemola and Devyani Sharma, eds. *The Oxford Handbook of World Englishes*. Oxford: Oxford University Press, 103-122.
- Mesthrie, Rajend. 2006. World Englishes and the multilingual history of English. *World Englishes* 25: 381–90.
- Mukherjee, Joybrato and Stephan Thomas Gries. 2009. Collostructional nativisation in New Englishes: Verb-construction associations in the International Corpus of English. *English World-Wide* 30(1): 27-51.
- Övergaard, Gerd. 1995. *The Mandative Subjunctive in American and British English in the 20th Century*. Uppsala: Almqvist and Wiksell.
- Peters, Pam. 1998. The survival of the subjunctive: Evidence of its use in Australia and elsewhere. *English World-Wide* 19: 87-103.
- Peters, Pam. 2009. The mandative subjunctive in spoken English. In Pam Peters, Peter Collins and Adam Smith, eds. *Comparative Studies in Australian and New Zealand English: Grammar and Beyond*. Amsterdam: Benjamins, 125-137.
- Quirk, Randolph, Sidney Greenbaum, Geoffrey Leech and Jan Svartvik. 1972. *A Grammar of Contemporary English*. London: Longman.
- Ronan, Patricia. 2016. Language relations in early Ireland. In Raymond Hickey, ed. *Sociolinguistics in Ireland*. Cambridge: Cambridge University Press, 133-153.
- Sayder, Stefan. 1989. The subjunctive in Indian, British and American English: A corpus-based study. *Linguistische Arbeiten* 69: 58-66.
- Schneider, Edgar W. 2000. Corpus linguistics in the Asian context: Exemplary analyses of the Kolhapur Corpus of Indian English. In Maria Lourdes Bautista, Teodoro A. Llamzon and Bonifacio P. Sibayan, eds. *Parangal nang Brother Andrew: Festschrift for Andrew Gonzalez*, Manila: Linguistic Society of the Philippines, 115-37.
- Schneider, Edgar W. 2005. The subjunctive in Philippine English. In Danilo T. Dayag and Stephen Quackenbush, eds. *Linguistics and Language Education in the Philippines and Beyond: A Festschrift in Honor for Ma. Lourdes S. Bautista*. Manila: Linguistic Society of the Philippines, 27-40.
- Schneider, Edgar W. 2007. *Postcolonial English. Varieties Around the World*. Cambridge: Cambridge University Press.
- Schneider, Edgar W. 2011. The subjunctive in Philippine English – An updated assessment. In M. L. S. Bautista, ed. *Studies in Philippine English: Exploring the Philippine Component of the International Corpus of English*. Mandaluyong City: Anvil Publishing, 159–173.
- Schneider, Edgar W. 2017. Models of English in the world. In Markku Filppula, Juhani Klemola and Devyani Sharma, eds. *The Oxford Handbook of World Englishes*. Oxford: Oxford University Press, 35-57.

- Sedlatschek, Andreas. 2009. *Contemporary Indian English. Variation and Change*. Amsterdam: Benjamins.
- Strevens, Peter. 1992. English as an international language: directions in the 1990s. In Braj. B. Kachru, ed. *The Other Tongue. English Across Cultures*. (Second Edition). Urbana and Chicago: University of Illinois Press, 27-47.
- Strobl, Carolin, Torsten Hothorn & Achim Zeileis. 2009. Party on! A new, conditional variable-important measure for random forests available in the party package. *The R Journal* 1(2): 14–17.
- Strobl, Carolin, James Malley & Gerhard Tutz. 2009. Introduction to recursive partitioning: Rationale, application, and characteristics of classification and regression trees, bagging and random forests. *Psychological Methods* 14(4): 323-348.
- Szmrecsanyi, Benedikt, Jason Grafmiller, Benedikt Heller & Melanie Röthlisberger. 2016. Around the world in three alternations: modelling syntactic variation in varieties of English. *English World-Wide* 37(2), 109–137.
- Tagliamonte, Sali & Harald Baayen. 2012. Models, forests, and trees of York English: *Was/were* variation as a case study for statistical practice. *Language Variation and Change* 24, 135–178.
- Turner, John F. 1980. The marked subjunctive in contemporary English. *Studia Neophilologica* 52: 271-277.
- Tan, Ying-Ying. 2015. The Americanization of the phonology of Asian Englishes: evidence from Singapore. In Gerhard Leitner, Azirah Hashim and Hans-Georg Wolf, eds. *Communicating with Asia. The Future of English as a Global Language*. Cambridge: Cambridge University Press, 120-134.
- van Rooy, Bertus. 2010. Societal and linguistic perspectives on variability in World Englishes. *World Englishes* 29(1): 3-20.
- Vine, Bernadette. 1999. Americanisms in the New Zealand English lexicon. *World Englishes* 18(1), 13-22.
- Waller, Tim. 2017. *Studies of the Subjunctive in Present-Day English: A Critical Analysis of Recent Research, Leading to a New Diachronic Investigation of the Mandative Subjunctive*. Unpublished PhD Dissertation, University College London.