Mandative subjunctive vs. *should* in world Englishes:
A new take on an old alternation

Sandra C. Deshors
Michigan State University

Stefan Th. Gries
University of California, Santa Barbara
& Justus Liebig University Giessen

Abstract
This study explores the alternation between the mandative subjunctive and its modal alternative with *should* across native and non-native Englishes. Methodologically, we try to improve on existing standards by investigating over 3,300 occurrences of the alternation from the Corpus of Web-based Global English and annotated for a range of linguistic factors analyzed with a forest of conditional inference trees; also, we are exemplifying a new strategy for the use of random or conditional inference forests in corpus-based alternation studies. We obtain a forest with a significant prediction accuracies and a good *C*-score and discuss the strongest predictors of the subjunctive vs. *should* alternation across Englishes. Contrasting with existing research, our multifactorial results (i) suggest that in British English the mandative subjunctive may not be dying out as much as we thought and (ii) individual suasive verbs influence speakers' uses of the two variants more than their English variety.

Keywords: mandative subjunctive, *should* construction, syntactic alternation, world Englishes, random/conditional inference forests

1 Introduction

In the past few years, syntactic alternations such as the dative and the genitive alternations have attracted much scholarly attention both with regards to native and non-native language varieties. Theoretically, the majority of alternation studies assumes a usage-based model of language use and acquisition which generally posits that language acquisition, processing, and change are all influenced by (frequencies) of use/occurrence and co-occurrence being processed by largely domain-general cognitive processes. In the context of syntactic alternations, adopting such a usage-based theoretical framework with its emphasis on (co-)occurrence means that syntactic variants are contrasted on the basis of the linguistic contexts in which they occur. Methodologically, this body of research consists of empirical studies that, often, have adopted (multifactorial) corpus-based approaches to pinpoint the contextual linguistic features that set syntactic variants apart and to assess to what extent the co-occurrence of contextual linguistic features influence speakers' constructional choices.

However, one particular alternation that has not yet received much multifactorial attention is the mandative subjunctive (e.g. *He demanded this be done immediately*) vs. *should* (e.g. *He ordered the culprit should be punished*) alternation (MSvS). While the mandative subjunctive (MS) has alternatives that involve modals other than *should*, "[t]hat with *should* […] is its most direct rival in that there is no appreciable semantic difference between them" (Collins 2015: 25). Further, as noted in Turner (1980: 276), "the present subjunctive in *that* clauses remains a productive means of expression in Modern English and one which deserves more discerning attention than has previously been the case in grammars purporting to account for language use".

Since the late 1970's-early 1980's the MSvS alternation has been a relatively highly debated topic in dialectology. Although Peters (1998: 101) notes that "[f]or the users of standard English in Australia, USA and Britain (and New Zealand), subjunctives are evidently a continuing resource in the articulation of certain kinds of subordinate clause", highly contrasting patterns have been observed in British and American Englishes. For instance, Algeo (1992: 612) reports that "[e]ven "verbally impoverished" and "semiliterate" students of remedial English use the mandative subjunctive because it is the norm in American English". More specifically, "American prefers the present (or mandative) subjunctive in the subordinate clause of such constructions, whereas British prefers the modal 'should' and can use the indicative" (Algeo 1992: 600). Very recently, however, interest in the alternation has started to extend to the realm of world Englishes and has particularly caught the attention of scholars whose main focus lies on how ESL speakers handle the alternation in their non-native language. While Hundt (2018) is a very welcome first attempt at a multifactorial research design to contrast the two constructions across a range of English varieties, her results remain to be validated and we will outline ways in which her methodology can be fine-tuned. Thus, the present large-scale corpus-based study adopts a multifactorial statistical approach to investigate the MSvS alternation across American, Australian, British and Indian Englishes. With this approach, we are able to unveil nuanced usage patterns that would have otherwise been hard to identify.

In what follows, we contextualize our study (Section 2) by presenting what is known about the alternation we are investigating and its determinants (Section 2.1). Then, we situate MS and *should* constructions in the context of (non-native) English varieties by discussing the diversity and patterning of uses of the two constructions across Englishes (Section 2.2). Finally, we discuss the exploration of MS vs. *should* constructions from a methodological standpoint and present main existing limitations that have remained unaddressed (Section 2.3). In Section 3, we present our corpus approach and discuss our data-processing and analysis strategies; in Section 4, we present the results of our analysis, whose implications are discussed in Section 5.

2 Setting the stage

2.1 Mandative subjunctive vs. should *constructions*: what we know about the alternation

In English, the MS serves the specific function of conveying directives including, for example, commands, orders or requests (Hoffmann 1997), as illustrated in (1).

(1) I demand that this *be* made available to the public again (GloWbE, g12)

As we can see in the above example, similarly to the imperative, the MS, appearing in italics, is formed by using the base form of a verb (e.g. *be* in our example), and therefore can only be distinguished in the third person singular. Syntactically, it tends to occur in object complement clauses (e.g. *that this be made available to the public again*) following a suasive verb such as *demand*, *order*, *request* or *ask* among others. Although MSs can also be used after adjectives and nouns expressing an emotion, in the context of the present study, we will limit our discussion and analyses to complement clause structures. As Hoffmann (1997: 7) explains, there are two alternants to the MS: a periphrastic form with a modal verb (illustrated in (2)) and the indicative (illustrated in (3))

(2) He ordered that the culprits *should* be punished
(3) I insist that she *arrives* on time

While there is no strict consensus about whether *mandative subjunctive* should refer only to non-inflected subjunctives or whether it should also include its other variants, we are following Hoffmann (1997) and Hundt (1998, 2018) in including only modals as in (2) to make our results more comparable to theirs.

Existing research shows that the alternation between the MS and *should* constructions is clearly not random and that a number of factors co-determine the alternants' usage patterns. For instance, Hoffmann (1997) singles out semantics as one of those aspects by pointing out how subtle differences in meaning can play a major role in the choice of variant, but Kastronic & Poplack (2014: 72) suggest that the semantic of the two constructions *per se* may not be enough to pinpoint what triggers the alternation of the two constructions given that "the meanings typically associated with the subjunctive are (fittingly enough) modal, pertaining to the desires, fears, emotions or hopes of the speaker or subject", thus making it harder to discern the forces at play behind the alternation. Correspondingly, Hoffmann (1997: 41-42) recommends that "other features such as syntactic and semantic constraints and perspective must also be taken into consideration".

Accordingly, several contextual linguistic features influencing the alternation have been identified. For instance, Hoffmann (1997) and Hundt (2018), among others, have shown that considerable differences exist between how much different main-clause suasive verbs attract MSs. According to Hoffmann (1997), *demand*, *order*, and *request* prefer the mandative subjunctive (some, such as *demand*, strongly), particularly with a non-inflected subjunctive, whereas *propose* prefers the modal variant; thus Hoffmann (1997: 26) concludes that "analysing mandative sentences as a unified grammatical phenomenon makes little sense [as the] differences between the individual suasive items are simply too large for such an undertaking".

Beyond suasive verbs, Algeo (1992: 600) explains that the choice of a superordinate governing expression may be involved in the choice of option in mandative constructions. More specifically, the presence/absence of subordination and particularly the presence/absence of a *that* complementizer introducing the mandative subjunctive or *should* constructions are important factors in understanding the alternation (Johansson 1979, Hoffmann 1997, Kastronic & Poplack 2014, Hundt 2018). According to Kastronic & Poplack (2014: 72), "the subjunctive variant is only admissible under specific subjunctive triggers when these occur in a legal subjunctive-selective context (introducing a subordinate clause headed by *that*)". What is particularly interesting, however, is how the presence/absence of the complementizer and its potential influence on the choice of mandative construction seems connected to the grammatical subject in the matrix clause. As noted in Hoffmann (1997: 61-62), according to Elsness (1984), zero connective is much more frequent if the matrix involves 1st or 2nd person subjects. When that is the case, the link between the two clauses is felt to be especially close.

Finally, voice and negation in the subordinate clause have also been investigated. Both Turner (1980) and Hornoiu (2015) observe that the (mandative) subjunctive is correlated with the passive voice. While negation is known to play a part in the use of the subjunctive in present-day English (Waller 2017), it tends to occur relatively rarely with mandative sentences (Hoffmann 1997). However, whether or not negation influences speakers' choice of one construction over the other when it *does* occur still remains to be established.

In sum, the factors most prominently discussed in previous work can be summarized as in Table 1. In the table, the center column includes the variables the literature has identified as

influencing the alternation we explore, the left column includes the levels of each variable that have been associated with the modal construction and the right column includes the levels of each variable that have been associated with the subjunctive construction.

Table 1 Overview of variables affecting the distribution of mandative subjunctives and *should* constructions

Modal construction example: *Picard demanded Cmdr Data should not be punished.*	Feature/predictor	Subjunctive example: *Picard demanded Cmdr. Data not be punished.*
propose	LEMMAMATRIX (Hundt 2018)	*require*, *request*, *demand*
-	LINKAGE (Kastronic & Poplack 2014)	*that*
	VOICE (in subordinate clause)	passive (Turner 1980; Hornoiu 2015); active (Hundt 1998)
presence	NEGATION (in subordinate clause; Hoffmann 1997)	
British	English VARIETY	American
	PERSONSUBJECT (Hundt 2018)	third person subject

Moving forward, the above-described body of research points towards the need to not only continue to account for contextual linguistic factors but, crucially, to account for those in a more integrated fashion: which factors have a predictive impact at the same time and which factors interact with each other e.g. by reinforcing or weakening each other in the context of all other factors?

2.2 The relevance of the mandative subjunctive vs. should *alternation for world Englishes*

Existing monofactorial literature shows that the MSvS alternation behaves very differently depending on the native English variety in which it occurs. For instance, based on diachronic research, Hornoiu (2015: 3) observes two main trends in the development of MSs that distinguish British and American Englishes: first, that the use of periphrastic constructions with *should* is less frequent in American than in British English (in line with Hoffmann 1997 and Leech et al. 2009) and second, that American English has been found to be leading World Englishes in an MS revival.[1] This variation in the uses of MSvS constructions has led a number of scholars such as Hornoiu (2015) to question whether, in certain varieties such as British English, the MS is to some degree dying out (Hundt 1998: 171) while possibly experiencing revival in other varieties such as American (Kastronic & Poplack 2004) and Australian English (Peters 1998).[2] According to Collins (2015: 17), from a historical perspective, this MS revival can be considered a postcolonial 'revival', given the steady decline of the previously-productive mandative had suffered from Early Modern English until the 19th century. Most importantly given our purposes, however, this possible revival of the MS is not only observed in American English but it is considered "American-led" (*ibid.*). What this suggests – in line with Peters (2015) – is that the revival of the MS in American English has started to infiltrate, so to speak, other native varieties, particularly Australian English

[1] See Algeo (1988: 20) for specific contrasting examples of the alternation across British and American Englishes.

[2] That being said, Hundt (1998: 171) finds some evidence that in British English the mandative subjunctive is losing some of its former stylistic connotations and that in the process "subjunctives are used in a wider-range of written text-types, they occur more frequently in the active voice today than thirty years ago and the co-occurrence of subjunctives and *that*-omission has increased".

followed by New Zealand English (Boberg 2004) where "the subjunctive construction is also common but "still in the process of revving up" (Hundt 1998: 171). This infiltration process is well described by Collins (2015):

> [i]n the revival of the mandative subjunctive […] AusE [Australian English] […] seems to be following the lead of AmE (which has maintained a preference for the subjunctive over *should* since the latter half of the 19th century) and to be eschewing the more conservative behavior of British speakers (who have maintained a dispreference for the mandative over the same period) (Collins 2015: 26)

However, although Övergaard (1995) shows that British MS usage has grown considerably since the 1960's (despite its preference for periphrastic *should* constructions) , Peters (1998: 89) warns that "[b]ecause Australian English shows influences from both British and American, it might reflect the current subjunctive habits of either" (see also Peters 2009). Indeed, Peters (2009) reports that the frequency of the MS in AusE overtakes that in BrE while approximating those recorded for AmE in other studies. It is necessary, however, to confirm these promising findings using state-of-the-art corpus and statistical methodologies.

Despite empirical evidence supporting the usefulness of exploring the MSvS alternation within world Englishes, only few studies have ventured beyond the circle of native varieties. Hundt's (2018) study is the first of its kind to conduct a relatively large-scale (approximately 1800 occurrences of the two constructional variants) multifactorial analysis of the alternation. Importantly however, the contribution consists of two separate and somewhat unrelated analyses of different (numbers of) English varieties: a first study using data from the International Corpus of English (ICE) and a second one using the Corpus of Web-based Global English (GloWbE). However, for the first time, both studies include some of the linguistic contextual features discussed in Section 2.1 as predictor variables, yielding results of a kind we have so far been missing. In a nutshell, and mainly based on the ICE data, Hundt (2018) observes that "variation in mandative sentences cuts across ENL [native English], ESL [English as a Second Language] and ESD [English as a Second Dialect] varieties" (Hundt 2018: 238), thereby justifying the expansion from existing research to non-native Englishes. More specifically, she finds that "IndE aligns closer to BrE than to SingE, another ESL variety, for instance". In the GloWbE study, she identifies factors that significantly influence constructional choice, namely suasive/trigger verb, variety, person subject and lexical verb in the subordinate clause. For instance, she observes that *request*, *require* and *demand* most strongly prefer a mandative subjunctive in the subordinate clause. However, at this point and despite the promise of her results, Hundt's (2018) GloWbE study also leaves open a variety of desiderata (described in more detail below) that we try to build on and improve on in the present paper with a view to paint a more precise picture of how second-language English differs from its native counterparts.

2.3 Exploring the alternation across Englishes: brief methodological insights and existing limitations

From a methodological perspective, exploring the MSvS alternation is less straightforward than it could appear. In fact, both Kastronic & Poplack (2004) and Waller (2017) have pointed out problematic methodological discrepancies between studies. Specifically, these scholars denounce the "disparities in both the number and identity of [subjunctive] triggers ranging from over 100 (e.g. Crawford 2009) to only four (e.g. Nichols 1987)" (Kastronic & Poplack (2004: 78)). Interestingly,

some notable disparity can be observed within Hundt's (2018) study itself where the two case studies in that single paper explore the same syntactic alternation in different corpora (ICE and GloWbE) but, curiously, with different subjunctive-triggering factors and contexts of use.

With regard to English varieties, the difference between her two studies is striking: her ICE study includes five ENLs (Canadian, British, Irish, New Zealand and Australian), four ESLs (Hong Kong, Indian, Philippine, Singaporean) and one ESD (Jamaican English)) whereas her GloWbE study is restricted to three varieties (two ENLs: American and British Englishes and one ESL: Indian English).[3] With regard to linguistic factors, the ICE study includes English variety, medium/register, trigger type (i.e. whether the suasive trigger is a lexical verb or an adjective), controlling subject (i.e. whether the grammatical subject in the matrix verb is a third or non-third person), verb in the subordinate clause (i.e. whether the lexical verb in the subordinate clause is *be* or any other lexical verb), negation and subordination (i.e. whether the subordinate clause is introduced overtly with a *that*-complementizer or covertly with a zero complementizer). By contrast, the GloWbE study excludes contexts with a zero-complementizer. Further, even though the factors controlling subject and verb in the subordinate clause were included in both studies, they were operationalized in the above-described coarse-grained fashion. Despite the importance of the predictor voice in the literature, this factor was excluded from both studies altogether. Further, although the ICE study includes eleven verbs, the GloWbE study is limited to a mere six suasive verbs. In addition, interactions between predictors are not accounted for although not only it is extremely rare for any alternation phenomenon to *not* involve such interactions but also, this omission is problematic as those interactions may play a significant part in how the two syntactic constructions alternate. Finally, her data sampling may not be ideal because while, for ICE, the specifics of the sampling are not all that clear, for GloWbE, it involves sampling the same number of hits (variable contexts) – 100 – for each verb. While that may seem appealing in how it guarantees the same number of data points per verb, this does also mean that the sample is not representative of the larger (language) population because the sampled verbs are of course not equally frequent in the language as a whole.

In sum, even though Hundt (2018) is an important first step in exploring MSs in world Englishes, the studies' research design – corpus methodology as well as statistical approach – require adjustments. In the present paper, we therefore revisit Hundt's GloWbE study and explain how several improvements are possible and, ultimately, required. Specifically, the present study is set up to address the following research goals:

- build on Hundt's multifactorial GloWbE study by adding factors reported to influence the MSvS alternation and exploring the potential effects of their mutual interaction on the alternation;
- proceed on the basis of a sampling scheme of trigger verbs that reflects the verbs' overall frequencies in the data;
- revisit Hoffmann's (1997: 26) claim that differences between the individual suasive items are simply too large to investigate mandative sentences reliably as a unified grammatical phenomenon and assess its validity across multiple English varieties by
    - using a better and more comprehensive statistical analysis of the data, namely one that (i) avoids making inferences about a random/conditional inference forest with a single tree (a practice that has been more widely adopted since Tagliamonte & Baayen 2012, but which we think is very problematic, see below), and that (ii)

[3] Hundt's (2018) multifactorial model is described in more technical terms in Section 3.2.

features what in a regression-modeling context would be interactions of predictors.

3 Methodology

3.1 Corpus data, data extraction and annotation

Our data were extracted from the Corpus of Web-Based Global English (GloWbE). Recently released, GloWbE is a 1.9 billion-word corpus of written English from twenty different countries.[4] The data exclusively consist of web-based material (e.g. newspapers, magazines, company websites, etc.), which is a written genre that has hardly been explored in the context of the mandative subjunctive (with the exception of Hundt 2018). Further, because of the sheer size of the corpus, GloWbE offers an unprecedented opportunity to explore more reliably than ever before mandative subjunctives which, as discussed earlier in Section 2.2 are less frequent in certain English dialects compared to others.

Regarding data extraction, we followed Waller's (2017: 204) recommendation to search the corpus for a list of subjunctive-triggering factors and then check the resulting concordances. Accordingly, we used R to extract all instances of the lexical trigger verbs *recommend*, *demand*, *require*, *suggest*, *propose*, *insist*, *request*, *ask*, and *order*, which were selected based on Hoffmann's (1997) list of suasive verbs; specifically, we selected the nine most frequent verbs with subjunctives that accounted for over 77% of all subjunctives. Altogether, a total of ≈1.376m occurrences of the 9 trigger verbs (and their linguistic contexts of use) were extracted from the American, Australian, British and Indian subsections of GloWbE. Given that extremely large amount of extracted occurrences, we decided to create a proportional sample that respected the marginal frequencies of varieties and verbs but also tried to minimize issues of data sparsity: Each variety was represented with a number of alternant data points that was proportional to the number of suasive verb hits in GloWbE and each verb was represented proportionally to its frequency in GloWbE but, for modeling purposes, with a minimum frequency of 22. Each extracted occurrence was then checked manually for whether it constituted an alternant until the minimum of relevant uses per lexical verb was reached.[5] Ultimately, a total of 3,343 occurrences of mandative subjunctive and *should* constructions were included in the study, annotated as described below and analyzed statistically. Table *2* presents an overview of the mandative subjunctive and *should* constructions included in the current study across the four English varieties in focus: British (BrE), American (AmE), Australian (AusE) and Indian (IndE) Englishes and Table *3* presents an overview of the number of occurrences of individual lexical verbs across these English varieties.

Once checked for syntactic relevance, all extracted constructions were annotated for the seven linguistic factors discussed above and listed in Table *4*.

[4] The countries included in the GloWbE corpus are the following: United States, Canada, Great Britain, Ireland, Australia, New Zealand, India, Sri Lanka, Pakistan, Bangladesh, Singapore, Malaysia, Philippines, Hong Kong, South Africa, Nigeria, Ghana, Kenya, Tanzania and Jamaica (https://21centurytext.wordpress.com/introducing-the-1-9-billion-word-global-web-based-english-corpus-glowbe/).

[5] In the absence of a formal distinction between a mandative and an indicative construction without a third-person subject, we relied on intuition to decide whether to include or exclude an occurrence from the study. As explained and illustrated in Hoffmann's (1997: 10), mandative constructions can be clearly distinguished from indicative constructions with non-third person subjects (see Hoffmann's 1997: 10 for specific examples). However, ambiguous cases were excluded from the study.

Table 2 Overview of the distribution of the mandative subjunctive and *should* constructions included in the current study across British, American and Australian Englishes

Construction	Australian Engl.	British Engl.	Indian Engl.	American Engl.	**TOTAL**
Modal (*should*)	101	424	115	150	**790**
Mandative subjunctive	431	815	225	1082	**2553**
TOTAL	532	1239	340	1232	**3343**

Table 3 Overview of the number of occurrences of individual lexical verbs across English varieties

Verb	**AustrE**	**BrE**	**IndE**	**AmE**	**TOTAL**
ask	33	87	22	85	**227**
demand	89	216	51	204	**560**
insist	66	103	37	119	**325**
order	39	84	25	83	**231**
propose	52	133	35	137	**357**
recommend	83	214	56	211	**564**
request	44	87	22	84	**237**
require	70	171	49	169	**459**
suggest	56	144	43	140	**383**

Table 4 Overview of the variables included in the study

Variable	**Variable levels**
CONSTRUCTION (dep. variable)	*modal*, *subjunctive*
VARIETY	*Australian* (*aus*), *British* (*gb*), *Indian* (*ind*), *American* (*us*)
LEMMAMATRIX (lemma occurring in the matrix clause)	*ask*, *demand*, *insist*, *order*, *propose*, *recommend*, *request*, *require*, *suggest*
LINKAGE (presence vs. absence of complementizer)	*that*, *zero*
VOICE (in the subordinate clause)	*active*, *passive*
PERSONSUBJ (grammatical subject in the subordinate clause)	*first plural*, *first singular*, *second*, *third plural*, *third singular*, *nonfinite*[6]
NEGATION (in the subordinate clause)	*affirmative*, *negative*
LEMMASUB (lemma occurring in the subordinate clause)	*be*, *use*. *give*, *have*, *do*, …

3.2 Statistical evaluation

Given the highly unbalanced, complex, and Zipfian distribution of our data, we did not go with the perhaps most widely-used method for this kind of data: generalized linear mixed-effects

[6] We recognize that the inclusion of non-third person grammatical subjects in our study could be considered somewhat controversial as scholars such as Johansson and Norheim (1988), Peters (1998) and Övergaard (1995) do not all agree on whether non-distinct forms (i.e., non-inflected verb forms) should be included or excluded from quantitative analyses (Hundt 2018). However, in the present study, we opted to include these non-distinct forms, in line with Övergaard (1995). Our decision is based on Övergaard's (1995: 69) finding that "no non-inflected verb forms in the American corpora can be regarded as ambiguous as regards mood [and] judged from the same perspective as the British instances, "ambiguous" tokens are few".

modeling. As in Tagliamonte & Baayen, but also other more recent studies – Bernaisch, Gries, & Mukherjee (2014) or Deshors & Gries (2016) in English variety/learner research, Dilts (2013) or Matsuki, Kuperman, & Van Dyke (2016) for psycholinguistic applications – we opted for an approach based on random forests, an extension of classification and regression trees, here specifically the kind referred to as conditional inference trees (Hothorn et al. 2006); this is the same methodology as used by Hundt (2018), the only other truly multifactorial corpus study of MSvS. Random/conditional inference forests add additional layers of randomness to such an analysis: First, many different conditional inference trees are constructed on different bootstrapped samples of the data. Second, each split in a conditional inference tree is only permitted to choose from a randomly-chosen subset of the available predictors rather than all of them. The predictions of the forest then consist of amalgamating the multitude of trees that were generated and their 'votes' for the out-of-bag cases.[7] Typically, the user has to specify only two hyperparameters (i.e., parameters that are defined before a statistical analysis begins and affect how it is conducted): the number of (randomly-chosen) predictors that may be considered at each split of each tree (*mtry*) and the number of trees grown. (*ntree*)

Especially since Tagliamonte & Baayen (2012), there has been a growing number of studies, in particular also Hundt (2018), that use something like the following approach for multifactorial alternation data – in particular data that are not amenable to regression modeling: (i) perform a forest analysis on the data,; (ii) report variable importance scores from the forest to assess each predictor's importance to the alternation; (iii) use a single classification/conditional inference tree on the complete data to visualize the predictors' effects. In this study, we are not following this approach. This is for two main reasons that previous research has ignored. First, the practice of interpreting a forest – i.e. a set of often 500 or even many more trees on randomly resampled data with different predictors at every split – on the basis of a single tree on all the data with neither level of resampling is highly problematic and can lead to misinterpretation of the patterns in the data. Second, the way in which forests are often interpreted – variable importance scores and (only occasionally) partial dependency scores – can fail dramatically at representing the nature of the effects in the data faithfully in terms of over- or underestimated variable importance scores and how predictors interact with each other (especially in smaller data sets and data sets that involve correlated predictors). Space does not permit a more detailed discussion here, suffice it to say that trees and forests, which are supposed to be very good at detecting and visualizing interactions are not necessarily as good as they are widely believed to be (see Winham et al. 2012, Boulesteix et al. 2015; Wright, Ziegler, & König 2016; Gries, to appear, for more discussion and exemplification).

In order to address all these issues we follow Gries's (to appear) recommendations: After a preliminary exploration of the data, which led to us unfortunately having to discard the variables NEGATION and LEMMASUB (because of their extreme imbalances), the first step of our statistical analysis consisted of manually creating a number of new predictors that essentially represent all two-way interactions: LINKAGE:VOICESUB, LINKAGE:PERSONSUBJ, LINKAGE:LEMMAMATRIX, LINKAGE:VARIETY, VOICESUB:PERSONSUBJ, VOICESUB: LEMMAMATRIX, VOICESUB:VARIETY,

[7] During the first of the two stages adding layers of randomness discussed above, the random forest algorithm splits the data up into a training and a test sample for each tree. Since this is typically done using sampling with replacement, not all cases make it into the training sample and the cases that are *not* used for training (i.e., building the tree), are referred to as out-of-bag cases, and they function as "a built-in test sample for computing the prediction accuracy of that tree [, the advantage being that that] is a more realistic sample of the error rate that is to be expected in a new test sample" (Strobl, Malley, & Tutz 2009:335).

PERSONSUBJ:LEMMAMATRIX, PERSONSUBJ:VARIETY, and LEMMAMATRIX:VARIETY. These were then added as predictors to a forest of *ntree*=1500 conditional inference trees with the number of predictors eligible for each split set to *mtry*=5. We then evaluated the forest in two ways: First, we computed the forest's overall precision and recall (for predicting subjunctive over modals), its prediction accuracy, and its *C*-score to determine how well the forest identified structure in our data. Second, we computed the version of variable importance scores proposed in Janitza, Strobl, & Boulesteix (2013), which is neither based on Gini/impurity scores nor on error rates from categorical predictions but on the area under the curve (*AUC*), which (i) makes that measure not just rely on categorical predictions, but also uses the probabilistic strength of the predictions and (ii) puts the same weight on both levels of the response variable as opposed to error rates, which give more weight to the more frequent level of the response variable; in a case like ours, where the more frequent level of the response variable accounts for more than 76%, this is an important means to arrive at more instructive variable importance measures.

The next step was to determine which predictors' effects to discuss, because, unlike with significance tests in a regression model, we are not aware of an obvious or natural cut-off point that determines which predictors' variable importance scores are high enough to merit discussion and which are not (the only recommendation we have ever seen is merely a heuristic). We therefore approached this question using a global surrogate model on the forest's predicted probabilities of subjunctive use. A global surrogate model is a statistical model of a kind that is fairly easily interpretable (such as linear regressions) and which is used to make the output of a statistical model of a kind that is hard to interpret (such as neural nets, support vector machines, random/conditional inference forests, and other black-box-like algorithms) easier to comprehend. Note that we are not using the GSM to interpret the forest, we are doing something simpler: We are using the GSM solely as a diagnostic tool that allows us to decide which predictors of the forest to discuss. Specifically, we fitted a linear regression model such that

- the dependent variable was the (logit-transformed) predicted probability of subjunctives obtained from the forest;
- the eligible predictors were all predictors used in the forest;
- we used a forward-selection modeling process adding variables in the order of the *AUC*-based importance scores until the relative likelihood of the new model did not increase drastically anymore (in a way similar to the use of scree plots in factor analysis).

Thus, after this process, we had an inventory of all predictors from the forest that the GSM considered important for the forest's interpretation; these were then summarized and visualized by computing the average observed probabilities of the two construction; in this analysis, these are virtually indistinguishable from the average predicted probabilities that are usually reported in regression-based analyses (this is not a sign of problematic overfitting – i is a sign that the GSM was able to do what it is supposed to do, namely to identify how the forest arrived at its predictions).

4 Results

Overall, our results confirm the general truism that linguistic alternations are never truly monofactorial and that, therefore, a multifactorial approach is required. In the present case, such

an approach based on a conditional inference forest resulted in a good fit to the data. The out-of-bag prediction accuracy of the forest is 0.803, which is significantly better than the baselines of choosing the more frequent construction or choosing randomly ($p_{\text{binomial test}}<10^{-7}$ and $p_{\text{binomial test}}<10^{-93}$). The precision in predicting subjunctive (as opposed to modals) was 0.855, recall was 0.892. Both variable importance scores ranked the importance of the predictors in the same order. In regression-analytic parlance, 'main effects' such as LINKAGE, VOICE were all qualified within 'interactions' with LEMMAMATRIX and VARIETY, which is why we focus on the following four interactions here:

– LEMMAMATRIX : VARIETY ($importance_{\text{AUC}} = 0.0573$);
– LEMMAMATRIX : LINKAGE ($importance_{\text{AUC}} = 0.0528$);
– LEMMAMATRIX : VOICE ($importance_{\text{AUC}} = 0.0509$);
– LEMMAMATRIX : PERSONSUBJ ($importance_{\text{AUC}} = 0.0347$).[8]

In the global surrogate model, these four interaction predictors yielded an adjusted R^2 of 0.895 and adding another predictor would have only increased adjusted R^2 by 0.005, which is why we visualize and discuss only these effects here in order of importance. Each figure is a two-panel representation of an interaction of two predictors, one panel with a dotchart for each perspective (where *perspective* refers to which predictors is shown as nested into which other); the dotcharts are representing observed percentages of MS (as opposed to *should*) for combinations of predictors. Each panel also represents an overall observed baseline of MS uses (with a long vertical dashed line) and medians of group percentages (with short vertical dashed lines). In each panel of these figures, the levels of the predictors involved are sorted from top to bottom in increasing order of median MS frequency.

4.1 Interaction 1: LEMMAMATRIX and VARIETY

Although scholars have paid much attention to the MSvS alternation in order to understand how to distinguish better between English varieties, our analysis yields only one strong interaction involving VARIETY, the interaction between lemmas in matrix clauses and English variety, which suggests that VARIETY on its own may be less strong a discriminator than is believed. Considering the left panel of Figure 1, we can see what one might call an overall main effect of LEMMAMATRIX with a cline of verbs strongly preferring subjunctives (*ask* > *request* > *demand*) over verbs with a weaker preference for subjunctives (*require* ≈ *recommend* > *order*) to verbs with a preference for modals (*insist* < *propose* < *suggest*). In addition, there is what one might call the main effect of VARIETY averaging across verbs: The right panel of Figure 1 shows that BrE and IndE pattern together preferring modals more than AusE and AmE. However, Figure 1 also clearly instantiates the kind of results that motivated the inclusion of an interaction predictor in the first place: the verbs' preferences vary – sometimes enormously – across varieties in three different ways: (i) verbs that are primarily used with MSs across varieties (*ask* (but see below), *request*, *demand*, and *require*), (ii) one verb that is uniformly preferred with *should* construction across varieties (*suggest*), and (iii) verbs – some of which were not included in Hundt's Figure 8 (copied in Figure A1 in the appendix) analysis – that exhibit considerable variation across varieties (*propose*, *insist*, *order*, *recommend*, and, perhaps, *ask*, given its lower percentage of MSs in IndE).

[8] Note that the absolute values of variable importance scores are usually not interpreted; we are following Strobl et al.'s (2009:342) suggestion to "[rely] only on a descriptive ranking of the predictor variables."

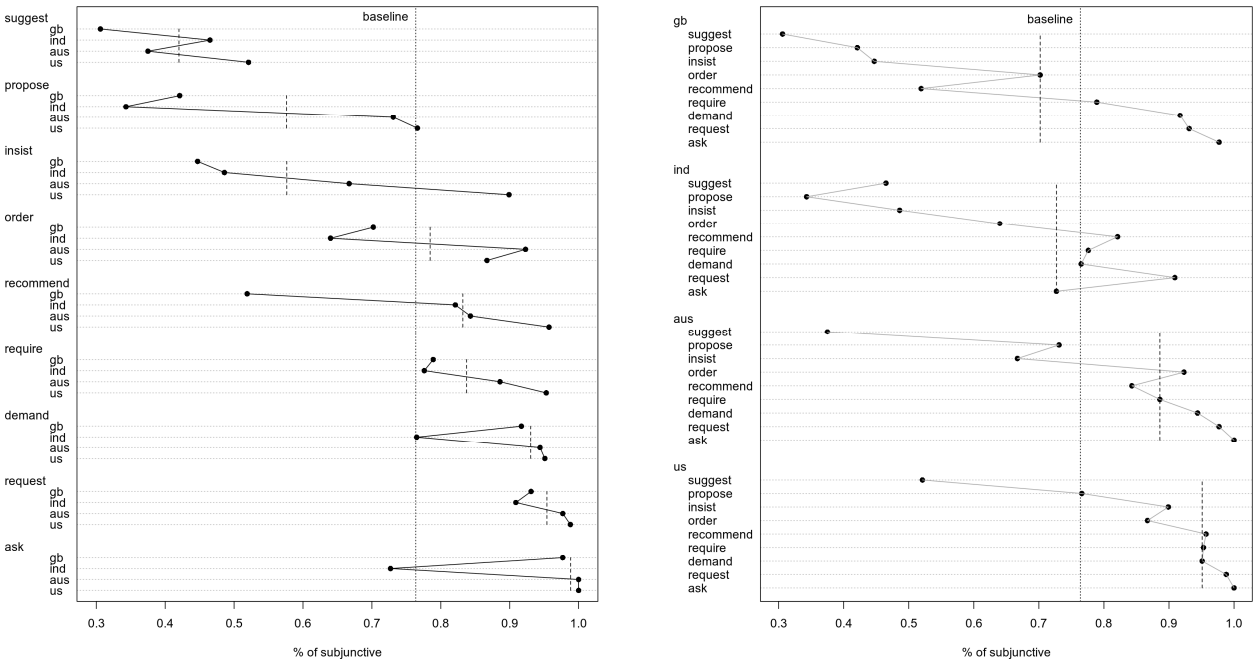

Figure 1 The interaction between LEMMAMATRIX and VARIETY

Let us compare our results to Hundt's (2018) more comprehensive GloWBe analysis in her Figure 8. In her data, *demand*, *request*, and *require* all strongly prefer subjunctives but to slightly different degrees: least so in IndE, more in BrE, most in AmE (with *demand* preferring subjunctives a bit less than *request* and *require*). We find a similar trend in that *demand*, *request*, and *require* also have the lowest subjunctive percentages in IndE compared to the native varieties, followed by BrE, and the AmE and AusE. Our results diverge from Hundt a bit such that, in our AmE data, *require* and *demand* pattern more alike than each does with *request*.

The other three verbs in Hundt's (2018) Figure 8 are *order*, *propose*, and *recommend*. On the whole, these also prefer subjunctives – the only configurations that actually predict *should* are, BrE and IndE, *propose* in general as well as *order* and *recommend* with non-third person and verbs other than *be*. Here, our results somewhat differ: while *propose* in BrE and IndE also strongly prefers *should* in our data (compared to AmE/AusE), *recommend* in BrE, but *not* IndE or anywhere else, strongly prefers *should*, and *order* prefers *should* more strongly in IndE than in BrE. Also, in Hundt's data, *order* and *recommend* are grouped together in both AmE and in BrE/IndE, whereas we find that, while their overall preference for subjunctives is similar, *recommend* in particular is quite diverse. Indeed, with regard to *order*, *propose* and *recommend*, while Hundt's classification tree indicates a split between AmE on the one hand and BrE and IndE on the other hand, our results suggest that amalgamating these latter varieties may be premature despite their common general trends.

With regard to verbs not included in Hundt's conditional inference tree, we find that *suggest* has an overall strong preference for modals whereas *insist* is, together with *recommend*, the verb exhibiting most marked differences between varieties: a strong preference for modals in BrE and IndE, an intermediate position in AusE, and a fairly strong preference for subjunctives in AmE.

4.2 *Interaction 2: LINKAGE and LEMMAMATRIX*

As mentioned earlier in this paper, the variable LINKAGE is one that, although identified in existing work as an important aspect of the MSvS alternation (see Hoffmann 1997), has so far not been part of a multifactorial analysis. As Figure 2 illustrates, the verbs' impact on the MSvS alternation is clearly not always the same: we can distinguish several different groups of verbs.

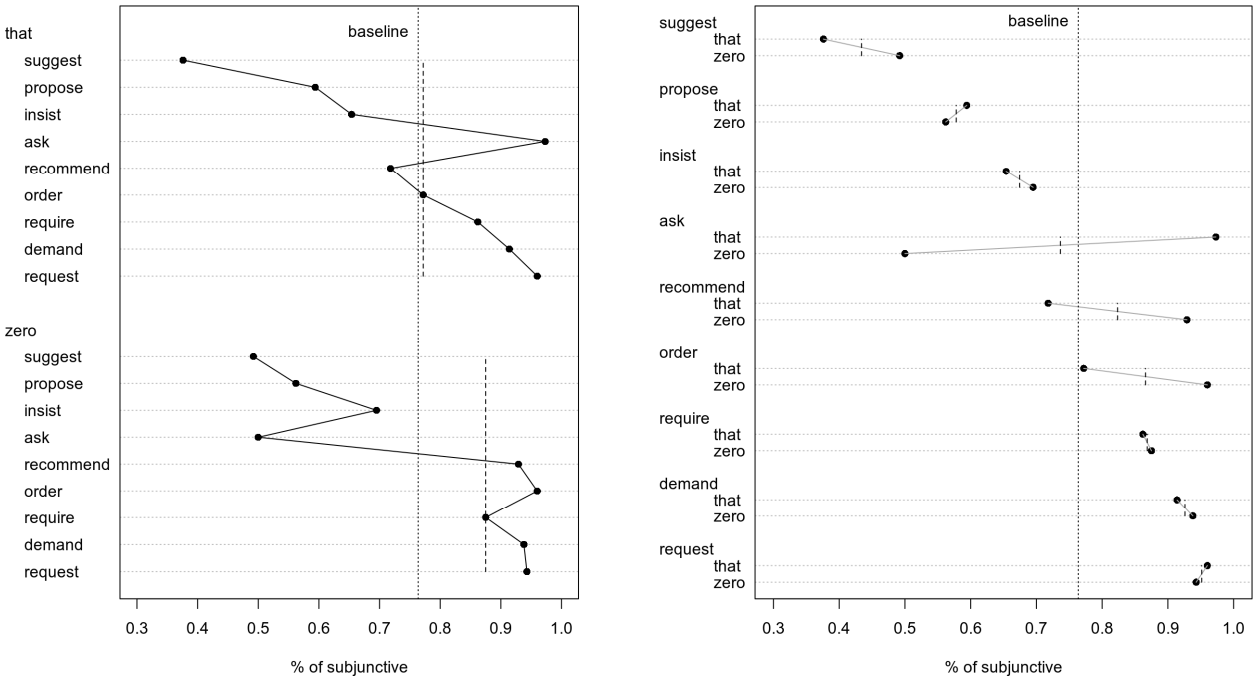

Figure 2 Interaction between LINKAGE and LEMMAMATRIX

First, *suggest*, *propose*, and *insist* all have a clear preference for *should* with only minor differences depending on LINKAGE. Second, *request*, *demand*, and *require* all have a clear preference for subjunctives with only minor differences depending on LINKAGE. Third, there is a third group of three verbs whose constructional preference differs depending on LINKAGE, i.e. where we find what in regression modeling is an interaction. These three verbs again come in two groups: one is *recommend* and *order*, which strongly prefer subjunctives when LINKAGE is zero, but whose preference for subjunctives decreases considerably in the presence of *that*. On the other hand, there is *ask*, which exhibits the largest interaction effect of all verbs: With no linking element, *ask* has a preference for *should* that is nearly as strong as that of *suggest* and even stronger than that of *propose* and *insist* – however, *ask* together with *that* has the strongest preference for subjunctives of all verbs.

4.3 *Interaction 3: VOICE and LEMMAMATRIX*

Similarly to LINKAGE, VOICE is a factor that has not yet been included in a multifactorial analysis of the alternation and it is neither a factor that has so far been studied on a large-scale basis across native and second-language English varieties. Figure 3, in particular the left panel, reveals that the overall preference for subjunctives is nearly the same for both actives and passives with a very slight stronger preference for subjunctives in passives, i.e. the main effect of VOICE is virtually non-existent; this effect is compatible with Algeo's (1992:607) observation that in passive

sentences the subjunctive is still the majority choice. However, as is more easily seen in the right panel, we again find a (weaker) interaction effect: While most verbs' preference is nearly identical regardless of VOICE (*suggest*, *insist*, *require*, *demand*, *ask*, and *request*), *propose* and *order*, on the other hand, have more subjunctives in the passives than in actives (although *propose* prefers *should* with both voices whereas *order* switches preference depending on VOICE), whereas *recommend* prefers subjunctives in actives and should in passives. Overall, these results are interesting for two main reasons: first, because for the first time they indicate that verb specificity needs to be considered in tandem with the effect that voice has on the MSsV alternation (we will return to the importance of verb specificity further below in our discussion section). Second, they reveal that the impact of voice on the MSvS is much more nuanced than previously believed. Specifically, they down tone previous research by Turner (1980) and Hornoiu (2015) that associated more categorically the subjunctive with the passive voice and research by Hundt (1998) that associated it more strongly with the active voice.

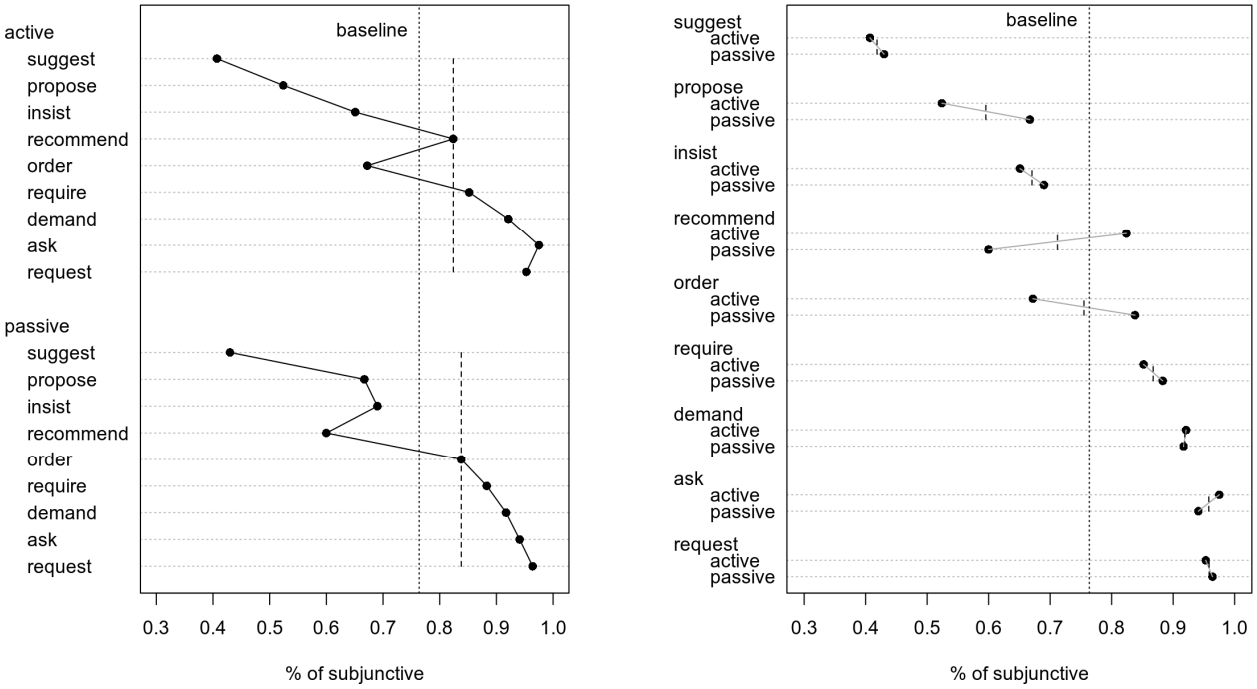

Figure 3 Interaction of VOICE and LEMMAMATRIX

4.4 Interaction 4: PERSONSUBJ and LEMMAMATRIX

Let us now move on to our last significant interaction, PERSONSUBJ and LEMMAMATRIX. Given the somewhat debated issue of the extent to which inflected forms should hold a place in an investigation of the MS (see note 6 in the current paper), the results in Figure 4 should of course be viewed as tentative and interpreted with a grain of salt. The most obvious finding from the right panel is the already discussed order of verbs in terms of subjunctive preference from least (*suggest*, then with some distance, *propose* and *insist*) to highest (*demand*, *order*, *ask*, and *request*). However, it is also clear that the verbs differ in terms of how much they interact with PERSONSUBJ: the preferences of *ask*, *demand*, *request*, and *require* do not vary much across different person/number combinations while the others verbs, and in particular *suggest* and *insist*, vary considerably across the levels of PERSONSUBJ.

The left panel is perhaps a bit harder to interpret: One can see that second person and first-person plural have the highest median percentages of subjunctives, with several verb forms exclusively occurring in the subjunctive. At the same time, even with these two levels of PERSONSUBJ, *suggest* and *propose*, which are strongly associated with lower levels of occurrence of subjunctives in general, lead to low percentages of subjunctives. Also, even the person-number combination with the lowest proportion of subjunctives – third person singular – still has an overall average number of subjunctives, showing that most person-number combinations really vary more because of the very strong effect of LEMMAMATRIX. The only level of PERSONSUBJ that exhibits less variability across verbs is first person singular, which is the only person where *suggest* does not strongly prefer *should*, where *request*'s preference for subjunctives is not close to 100% or at least above 90%, and where *order*'s preference for modals is higher than average.

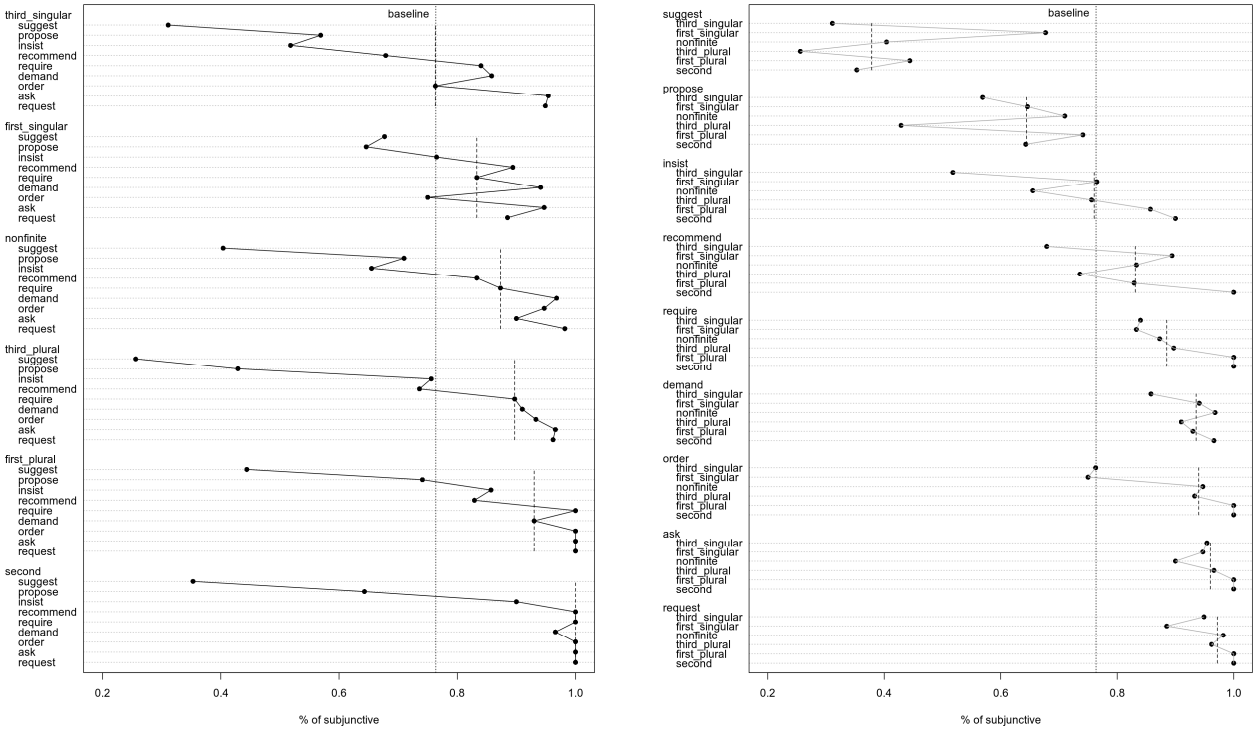

Figure 4 Interaction of PERSONSUBJ and LEMMAMATRIX

5 Discussion

With our analysis, we set out to revisit the already well-studied MSvS alternation from the perspective of multifactoriality. As the rapidly increasing body of research using multifactorial statistical methods demonstrates, this type of approach helps explore alternating linguistic constructions by providing sophisticated tools to explain why speakers choose one construction over another and why their constructional choices can vary systematically across Englishes. With regards to the MSvS alternation specifically, although our study follows Hundt (2018) as the second multifactorial analysis of the alternation, our methodological set up is the first one to (i) at least initially include all linguistic predictors known to affect the alternation with a proportional sampling scheme and (ii) assess, by means of a random forest analysis enhanced by interaction

predictors how the *combined* effects of these predictors influence speakers' constructional choices. These are important steps as they allowed us to capture, in the most realistic way possible yet, the complexity of the linguistic contexts in which the MS and *should* constructions are used. As a result of our methodological design, it emerges that long-debated issues central to the MSvS alternation, such as the gradual disappearance of the MS in British English and its American-led revival in other varieties of English, are, to some degree, put back into question. Further, for the first time, our results allow us to make a connection between individual suasive verbs and the diachronic development of the MS across Englishes. Although individual suasive verbs have been given a prominent place throughout the literature on the MSsV alternation, as far as we know, to date, variation in the uses of these individual suasive verbs has not been taken into account in the context of the disappearance and revival of the MS in BrE and AmE. In what follows, we discuss in more detail the implications of our results with a specific focus on the diachronic development of the MS. In addition, we discuss the methodological implications of accounting for predictor interactions in random forest analyses.

With regards to diachronic change, as we mentioned in Section 2.3, previous research points towards different developmental patterns of the MS construction across English varieties. Indeed, while scholars such as Hundt (1998: 171; amongst others) have claimed that the MS is currently *dying out* in certain varieties of English such as BrE, other scholars such as Kastronic & Poplack (2004) and Peters (1998) have argued for a revival of the construction in other varieties such as AmE and AusE. Undoubtedly, these claims call for diachronic data in order to be fully validated. However, our synchronic data can nonetheless provide a valuable snapshot of the development of the construction and raise questions as whether the MS is truly *dying out* in BrE. Indeed, our results suggest that overall existing studies may have overestimated the disappearance of the MS in that variety. More concretely, in the right panel of Figure 1 we observed a median of 70% subjunctive with the suasive verbs we investigated even within the BrE data and in the specific cases of *demand*, *request* and *ask* which subjunctives occur over 90% of the time. Although it is true that in BrE and IndE the proportion of subjunctives is lower compared to AmE and AusE, the overall proportion of subjunctives in BrE and IndE still remains relatively high. These results have an important implication: with such high medians of subjunctives in very recent web-based data, it is hard to claim that the MS is *dying out* in BrE. As a result, this claim should not only be toned down but also be made more specific in the sense that the MS is not equally dying out in all linguistic contexts. Indeed, based on the LEMMAMATRIX and VARIETY interaction, the statement of *dying out* cannot be made felicitously for suasive verbs in general.

Based on our interaction results, it is clear that some previous research has seriously underestimated the role suasive verbs play in the MSvS alternation (across Englishes), as is obvious from the combined/interaction effects of LEMMAMATRIX with four other predictors, namely VARIETY, LINKAGE, VOICE and PERSONSUBJ. This finding yields an important disconnect between the current study and existing work in how it points to the necessity that the MSvS alternation simply *has* to make sure verb variation is included in all analyses. Despite the fact that overall our result confirms Hoffmann's (1997: 26) claim that mandative sentences cannot be investigated as a unified grammatical phenomenon due to the large differences between the individual subjunctive-triggering suasive verbs, our multifactorial methodological design allowed us to establish with more precision (i) how pervasive the effect of suasive verbs really is (i.e., with which specific contextual linguistic factors individual verbs have to co-occur with in order to influence the constructions' alternation patterns?) and (ii) to what exact degree it does so. While this level of quantitative precision may seem trivial, it outperforms existing work that is still

lacking in recent publications such as Collins (2015: 26) who finds that his study "revealed *a certain amount* of lexical conditioning in the occurrence of the mandative subjunctive" (our emphasis).

Importantly, underestimating the impact of verb specificity on the MSvS alternation could have misled scholars in their understanding of the relatively recent development of the MS in AmE and other varieties that tend to follow the American lead. As just mentioned, our results clearly show the central role matrix verbs play for the alternation so while existing literature does not ignore these verbs, traditionally, they are not accounted for systematically in terms of *both* (i) their main effect on the alternation (rather than their observed frequency of occurrence with each alternating constructional variant) and (ii) their joint effect with another predictor on the alternation. This is a critical point as it puts into perspective much of the existing work on the debated American-led revival of the MS. As far as we are aware, this existing literature does not account multifactorially for verb specificity as a contributing factor of the potential revival process. Zooming in on the AmE and AusE varieties, which, as we previously discussed have been claimed to be under-going a revival of the MS (Collins 2015, Peters 2015) under the leadership of AmE, our results are overall compatible with Collins (2015) in that the verb frequencies in the MS in AusE are more similar to AmE than they are to BrE. Our results also confirm Peters (2009) in that in our data the frequency of the MS in AusE does overtake that in BrE while approximating those recorded for AmE. However, we do stress that these similarities (along with subsequent discussions on the revival of the MS) do need to be considered in relation to the frequencies of suasive verbs and their constructional preference(s).

Moving on to statistical methodology and methodological/statistical implications, the current study exemplifies a number of important issues discussed in Gries (to appear), with important implications for the use of random/conditional inference forests. While they are becoming used more frequently in corpus linguistics, this is not without risks. Their deceptive simplicity notwithstanding, just about every single aspect of forests is currently being lively discussed in bioinformatics journals: sampling of data (with or without replacement), splitting criteria (Gini vs. *p*-values), variable importance measures (error rate vs. permutation-based vs. *AUC* (the latter two conditional or unconditional)), variable selection, whether random/conditional inference forests can capture or detect interactions in the presence of correlated predictors, imbalanced response variables etc., all of which affect the (quality of the) results … What informed our approach here is that summarizing a forest with single tree on all data is highly problematic and that nearly all current work involving forests in corpus linguistics does not even consider the notion of interactions of predictors. We therefore tried to improve on existing work in the field by promoting and exemplifying three aspects: We (i) added interaction predictors to the forest (following recommendations in predictive modeling literature), (ii) used the forest to compute *AUC*-based variable importance scores (which are better at handling the class imbalance problem that corpus-based alternation data often exhibit), and (iii) we used a global surrogate model to determine which predictors of the forest merit discussion (something that variable importance scores do not do straightforwardly).

The results/advantages of this process are rather striking and especially the benefits of (i) can be very simply clarified on the basis of any of Figures 1 to 4: Not including interactions would mean that, for every predictor, we would only have the overall mean percentages indicated with the short dashed vertical lines in each figure, whereas we have often seen how individual verbs stray extremely widely from that overall main effect (recall how much *propose*'s behavior differ across varieties in the left panel of Figure 1 or *ask*'s interaction with VOICE in the right panel of

Figure 2). And even partial dependence scores for, say, LEMMAMATRIX would also merely correspond to something like the nine dashed lines (1/lemma) in the left panel of Figure 1, but would *not* usually include how, within each verb, the varieties differ. Our way of including interactions is therefore a simple but motivated way to explore things at a greater level of detail without which, for instance, the degree to which any predictor's effect is in fact mediated by the verb lemma of the matrix clause is unknowable. We hope that this strategy, or others like it,[9] is one that can move the field along by providing more informative results with a principled data exploration/modeling approach.

References

Algeo, J. 1992. 'British and American mandative constructions' in C. Blank (ed.) *Language and Civilization: A Concerted Profusion of Essays and Studies in Honour of Otto Hietsch*, pp. 599-617. Frankfurt: Peter Lang.

Bernaisch, T., St. Th. Gries, and J. Mukherjee. 2014. 'The dative alternation in South Asian English(es): Modelling predictors and predicting prototypes', *English World-Wide* 35(1), pp 7-31.

Boberg, C. 2004. 'English in Canada: Phonology' in B. Kortmann, E. Schneider, K. Burridge, R. Mesthrie and C. Upton (eds) *A Handbook of Varieties of English*, pp. 351-365. Berlin & Boston: De Gruyter.

Boulesteix, A.-L., S. Janitza, A. Hapfelmeier, K. Van Steen, & C. Strobl. 2015. Letter to the editor: On the term 'interaction' and related phrases in the literature on Random Forests. *Briefings in Bioinformatics* 16(2). 338-345.

Collins, P. 2015. 'Diachronic variation in the grammar of Australian English. Corpus-based explorations' in P. Peters, P. Collins and A. Smith (eds.) *Grammatical Change in English World-Wide*, pp. 15-42. Amsterdam & Philadelphia: John Benjamins.

Deshors, S. C. and St. Th. Gries. 2016. 'Profiling verb complementation constructions across New Englishes: A two-step random forests analysis to *ing* vs. *to* complements', *International Journal of Corpus Linguistics* 21(2), pp. 192-218.

Dilts, P. 2013. Modelling phonetic reduction in a corpus of spoken English using Random Forests and Mixed-Effects Regression. Unpublished Ph.D. dissertation, University of Alberta, Edmonton.

Elsness, J. 1984. 'That or zero? A look at the choice of object clause connective in a corpus of American English', *English Studies* 65(6), pp. 519-533.

Gries, St. Th. to appear. On classification trees and random forests in corpus linguistics: some words of caution and suggestions for improvement. *Corpus Linguistics and Linguistics Theory*.

Gries, St. Th, T. Bernaisch and B. Heller. 2018. 'A corpus-linguistic account of the history of the genitive alternation in Singapore English' in S. C. Deshors (ed.) *Modeling World Englishes: Assessing the interplay of emancipation and globalization of ESL varieties*, pp. 245-279. Amsterdam & Philadelphia: John Benjamins.

Hoffmann, S. 1997. 'Mandative Sentences. A Study of Variation on the basis of the British National Corpus', Lizentiats-Thesis, University of Zurich.

[9] Alternatives to the including-explicit-interactions approach used here are discussed and exemplified in Gries (to appear) and its companion file.

Hothorn, T., K. Hornik, and A. Zeileis. 2006. 'Unbiased recursive partitioning: a conditional inference framework', *Journal of Computational and Graphical Statistics* 15(3), pp. 651-674.

Hundt, M. 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 L. Hans, S. Klintborg, M. Levin and M. Estling (eds.) *The Major varieties of English*, pp. 159-175. Växjö: Växjö University.

Hundt, M. 2018. 'It is time that this (*should*) be studied across a broader range of Englishes: A global trip around mandative subjunctives' in Sandra C. Deshors (ed) *Modeling World Englishes: Assessing the interplay of emancipation and globalization of ESL varieties*, pp. 217-244. Amsterdam & Philadelphia: John Benjamins.

Hornoiu, D. 2015. 'The Subjunctive in Present-day English: Revival or Demise?', *Language and Literature Studies* 26(1), 67-77.

Johansson, S. 1979. 'American and British English grammar: An elicitation experiment', *English Studies* 60, pp. 195-215.

Johansson, S. and E. H. Norheim. 1988. 'The subjunctive in British and American English', *ICAME* Journal 12, pp. 27-36.

Janitza, S., C. Strobl and A-L Boulesteix. 2013. 'An AUC-based permutation variable importance measure for random forests'. *BMC Bioinformatics* 14, pp. 119.

Kastronic, L. and S. Poplack. 2014. 'The (north) American English mandative subjunctive in the 21st century: Revival or remnant?', *University of Pennsylvania Working Papers in Linguistics* 20(2), pp. 71-80.

Leech, G., M. Hundt, C. Mair, and N. Smith. 2009. *Change in Contemporary English: Grammatical Study*. Cambridge: CUP.

Matsuki, K., V. Kuperman, and J.A. Van Dyke. 2016. 'The Random Forests statistical technique: An examination of its value for the study of reading'. *Scientific Studies of Reading* 20, pp. 20-33.

Nichols, A. E. 1987. 'The suasive subjunctive: Alive and well in the Upper Midwest'. *American Speech* 62, pp. 140-153.

Övergaard, G. 1995. *The Mandative Subjunctive in American and British English in the 20th Century*. Stockholm: Almqvist & Wiksell.

Peters, P. 1998. 'The survival pf the subjunctive: Evidence of its use in Australia and elsewhere'. *English World-Wide* 19, pp. 87-103.

Peters, P. 2009. 'The mandative subjunctive in spoken English' in P. Peters, P. Collins and A. Smith (eds.) *Comparative studies in Australian and New Zealand English: Grammar and beyond*, pp. 125-137. Amsterdam & Philadelphia: John Benjamins.

Strobl, C., J. Malley, and G. Tutz. 2009. An introduction to recursive partitioning: rationale, application, and characteristics of classification and regression trees, bagging, and randomf forests. *Psychological Methods* 14(4), pp. 323-348.

Tagliamonte, S. A. and R. H. Baayen. 2012. 'Models, forests and trees of York English: *Was*/*were* variation as a case study for statistical practice', *Language Variation and Change* 24(2), pp. 135-178.

Turner, J. F. 1980. 'The marked subjunctive in contemporary English'. *Studia Neophilologica* 52, pp. 271-277.

Waller, T. 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*. PhD

Dissertation, University College London.
Winham , S.J., C. L Colby, R. R Freimuth, X. Wang, M. de Andrade, M. Huebner, and J. M Biernacka. 2012. SNP interaction detection with Random Forests in high-dimensional genetic data. *BMC Bioinformatics* 13(164).
Wright, M.N., A. Ziegler, & I. R. König. 2016. Do little interactions get lost in dark random forests? *BMC Bioinformatics* 17(145).

Appendix 1

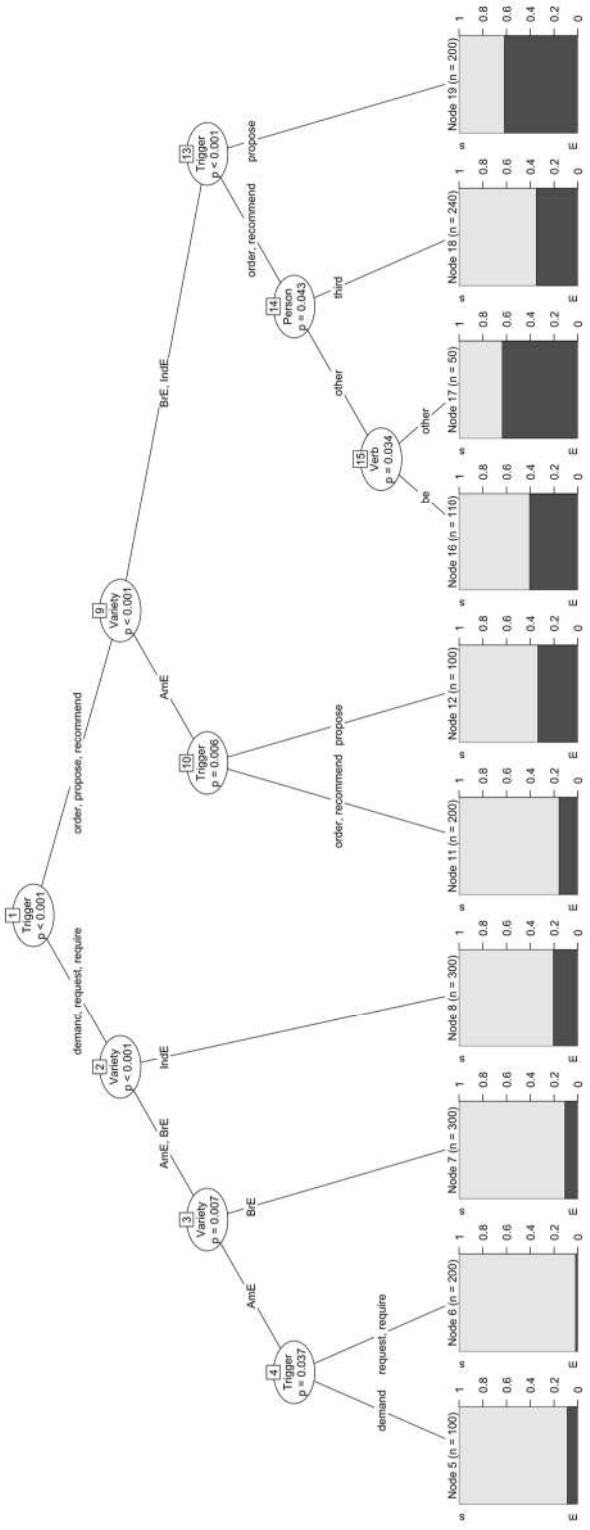

Figure A1 Hundt's (2018) Figure 8 showing variable importance (trigger and variety) in GloWbE data

Appendix 2: R code/results of the conditional forest analysis

```
# loading data from .csv file
summary(x <- read.table("06_data.csv", header=TRUE, sep="\t", quote="",
comment.char=""))
str(x)
'data.frame':     3343 obs. of  6 variables:
 $ VARIETY     : Factor w/ 4 levels "gb","us","aus",..: 1 1 3 1 1 1 2 ...
 $ CONSTRUCTION: Factor w/ 2 levels "modal","subj": 2 2 1 2 2 2 2 ...
 $ LINKAGE     : Factor w/ 2 levels "that","zero": 1 1 1 1 1 1 1 ...
 $ VOICESUB    : Factor w/ 2 levels "active","passive": 2 2 1 2 2 1 1 ...
 $ PERSONSUBJ  : Factor w/ 6 levels "first_plural",..: 3 3 6 6 6 6 1 ...
 $ LEMMAMATRIX : Factor w/ 9 levels "recommend","ask",..: 3 3 4 6 1 3 8 ...

# adding interactions
x <- cbind(x,
   x$LINKAGE:x$VOICESUB, x$LINKAGE:x$PERSONSUBJ, x$LINKAGE:x$LEMMAMATRIX,
   x$LINKAGE:x$VARIETY, x$VOICESUB:x$PERSONSUBJ, x$VOICESUB:x$LEMMAMATRIX,
   x$VOICESUB:x$VARIETY, x$PERSONSUBJ:x$LEMMAMATRIX, x$PERSONSUBJ:x$VARIETY,
   x$LEMMAMATRIX:x$VARIETY)
names(x)[7:16] <- c("LINVOI", "LINPER", "LINLEM", "LINVAR", "VOIPER",
   "VOILEM", "VOIVAR", "PERLEM", "PERVAR", "LEMVAR")

# fitting conditional inference forest
library(party)
set.seed(150270); rf.p.1 <- cforest(CONSTRUCTION ~
   LINKAGE+VOICESUB+PERSONSUBJ+LEMMAMATRIX+VARIETY+
   LINVOI+LINPER+LINLEM+LINVAR+VOIPER+VOILEM+VOIVAR+PERLEM+PERVAR+LEMVAR,
   data=x, controls=cforest_control(ntree=1500, mtry=5))

# compute AUC-based variable importance scores
sort(round(varimpAUC(rf.p.1), 4), decreasing=TRUE)
# LEMMAMATRIX      LEMVAR      LINLEM      VOILEM      PERLEM     VARIETY
#      0.0786      0.0573      0.0528      0.0509      0.0347      0.0256
# VOIVAR      LINVAR      PERVAR      VOIPER  PERSONSUBJ      LINPER
# 0.0177      0.0174      0.0170      0.0062      0.0050      0.0050
# LINVOI    VOICESUB     LINKAGE
# 0.0028      0.0017      0.0005
```