

The range of linguistic units: Distance effects in English mandative subjunctive constructions¹

THOMAS BERG

University of Hamburg

TIM ZINGLER

University of New Mexico

ARNE LOHMANN

University of Düsseldorf

(Received 6 February 2018; revised 13 March 2019)

This study examines the role of distance in the decision among grammatical variants. The empirical test case is the English mandative subjunctive construction, which co-occurs with an embedded modal auxiliary, a subjunctive or an indicative verb form. The fact that the subjunctive is triggered by specific lexical items allows one to measure the distance between the triggering unit and the target verb. This distance is found to play a significant role in the grammatical decision process. With increasing distance between trigger and target, the probability of selecting a modal auxiliary increases and the probability of selecting the subjunctive decreases. The theoretical account hinges on the range and strength of linguistic units. Syntactic units (i.e. modals) are claimed to have a wider range than morphological units (i.e. indicative and subjunctive). Furthermore, the indicative is claimed to have a wider range than the subjunctive. Varying ranges are interpreted as varying decay rates. The lower decay rate of syntactic as compared to morphological units results from the syntactic level being superordinated to the morphological level in language production. The inclusion of the semantic and the phonological levels confirms that the position of a level in the structural hierarchy determines its range.

KEYWORDS: activation, decay, distance, hierarchy, subjunctive, variation

[1] A foretaste of the final product was given in a pleasant atmosphere at the University of Bamberg in June 2016. Our heartfelt thanks go to Tayo Takada, Gonca Bakir and André Geisler who invested much effort in the data coding procedure, to Julia Schlüter for helpful comments and to Benedikt Szmrecsanyi for sharing some of his data with us. We would have liked to have seen Tayo's name at the top of the page, but unfortunately things do not always turn out the way they should. The final version has greatly benefited from the suggestions of the three *Journal of Linguistics* referees, to whom we gladly express our gratitude.

1. ON THE ROLE OF DISTANCE

One of the notable properties of language is the ubiquity of dependency relationships between linearly ordered units (e.g. Gildea & Temperley 2010, Futrell, Mahowald & Gibson 2015). An adjective may depend on a noun, an object NP on a verb and a subordinate clause on a main clause. The distinction between independent units (i.e. heads) and dependent units (i.e. modifiers or dependents) in the linear representation of speech prepares for the notion of distance. If distance is a relevant factor in linguistic analysis, it may be expected to influence the interaction of heads and dependents as follows: the smaller the distance between two given units, the more likely one is to impact on the other. The issue is not only whether or not an interaction takes place but also whether different distances may have different effects.

As a matter of fact, there is solid evidence to vindicate the role of distance (e.g. Gibson 2000, Pietsch 2005, Gillespie & Perlmutter 2011; see also Rosenbaum 1970 and Horn 1986 for early discussions of the Minimum Distance Principle). To set the stage, three related areas in which distance plays an important part will be briefly reviewed – agreement, grammatical variation and mood. What they have in common is that the head does not categorically determine the form of the dependent.

Subject–verb agreement in English provides a clear demonstration of the effect of distance. The critical case involves conflicting information from different potential controllers and their linear position relative to the target. Consider existential constructions in English, as illustrated in examples (1) and (2) from Morgan (1972: 281; for the same effect in other languages, see Plank (1985: 127) on German, Sridhar (1990: 246) on Kannada (Dravidian) and Jeschull (2004: 251) on Chechen (Nakh-Daghestanian)).

- (1) There was/*were a man and two women in the room.
- (2) There *was/were two women and a man in the room.

Except for the ordering within the logical subject, the two sentences are structurally identical. Whatever difference there is can therefore be put down to the opposite order of the nominal constituents in the coordinated NP. If the singular noun occurs next to the verb, singular agreement is the rule, as in (1). By contrast, if the plural controller is adjacent to the verb, plural agreement is observed, as in (2). It is thus the noun closest to the verb rather than the entire NP which is elected as controller. This fact assigns a role to distance.

Strictly speaking, the above pair of sentences illustrates a binary adjacency effect but remains silent on the possibility of gradient effects with varying distance. This issue will be addressed on the basis of Levin's (2001) analysis of collective nouns in English which allow both singular and plural agreement on the verb, as exemplified in (3) and (4) from Levin (1998: 106; 2001: 96).

- (3) The British royal family is now clear of haemophilia.
 (4) My family always get upset if I talk about it.

Levin's results show a remarkable interaction of distance and agreement type. There is a gradual decrease in singular agreement and a concomitant increase in plural agreement with an increasing number of words intervening between target and controller.

The sensitivity of agreement to the distance between target and controller is restricted neither to English nor to number agreement. Köpcke, Panther & Zubin (2009) capitalize upon a mismatch between grammatical and referential gender in some German nouns such as *das Mädchen* 'the girl', which is grammatically neuter but referentially feminine. Their analysis reveals that the rate of referential gender increases, and that of grammatical gender decreases, with increasing distance between target and controller. This observation ties in with data from French (Corbett 1991: 227), Cairene Arabic (Belnap 1999: Table 2) and Old English (Curzan 2003: 99).

The choice between alternative grammatical structures is also sensitive to distance. One of the major factors determining serial order is the short-before-long principle (e.g. Behaghel 1909, Hawkins 1994, Lohmann & Takada 2014). Hawkins's (1994) performance theory relies, among other things, on the distance between the heads of phrases. In head-initial languages, this distance is smaller when the shorter phrase precedes the longer one. The preference for short distances can be seen in many structures ranging from so-called particle movement (e.g. Gries 2003, Lohse, Hawkins & Wasow 2004) to so-called dative alternation (e.g. Arnold et al. 2000, Bresnan et al. 2007) and the order of PPs (e.g. Hawkins 1999, Wiechmann & Lohmann 2013). As a further domain, consider the contrast between pied-piping and preposition stranding, as exemplified in (5) and (6), respectively.

- (5) But questions with which committee members taunted bankers appearing as witnesses left little doubt that they will recommend passage of it.

(from Hoffmann 2011: 93)

- (6) But the questions which he taunted us with left little doubt. . .

(from Trotta 2000: 188)

The preposition *with* is stranded in (6) but pied-piped in (5). The choice between the two constructions is influenced by the amount of material coming between the relativizer and the stranded preposition. The greater the distance between the two critical items, the lesser the likelihood of preposition stranding (Trotta 2000, Hoffmann 2011). The same result was obtained for relative clauses in Brazilian Portuguese by Tarallo (1986) and for English interrogative clauses by Gries (2002).

A superficially different type of grammatical variation is the alternation between ‘something’ and ‘nothing’. Several structures in English display a contrast between a fully fledged and a reduced variant. The latter typically differs from the former in that there is no independent function word. This contrast is exemplified in (7) on the basis of the verb *to stop* which may or may not take the preposition *from* (Rohdenburg 1999).

- (7) (a) The police stopped the fans who had no valid tickets (from) entering the grounds.
 (b) The police stopped the fans (from) entering the grounds.
 (c) The police stopped them (from) entering the grounds.

The three examples differ in the complexity of the object NP of the matrix verb. A nominal object expanded by a relative clause is shown in (7a), a bare nominal object in (7b) and a pronominal object in (7c). Rohdenburg (1999) argues that the use of the preposition correlates with an increasing complexity of the object. However, the empirical effect can be captured equally well by invoking distance: the greater the distance between the verb and the slot for the preposition, the higher the likelihood of choosing the prepositional option.

The third and final area moves us closer to the empirical focus of the present paper. It is distantly related to agreement and documents an influence of the verb in the main clause on the verb in the subordinate clause. In French, the mood of the embedded verb may be determined by certain triggers in the matrix clause. Some triggers permit variation between indicative (IND) and subjunctive (SUB) verb forms in the subordinate clause, as illustrated in (8) and (9) from Poplack (1992: 250) and Poplack, Lealess & Dion (2013: 172), respectively.

- (8) Bien certain, faut qu'ils aient
 well certainly necessary that they have(SUB)
 une place eux- autres aussi pour vivre.
 a place they- others also for living
 ‘Well, of course, they should have a place to live, too.’
- (9) ...il faut au moins tu te c-
 it (is) necessary at least you yourself
 tu fais pas mal à l'autre personne.
 you do(IND) not wrong to the other person
 ‘... you have to at least you – you do not hurt your opponent.’

In the citation form, the impersonal verb *falloir* takes the form *il faut que* ‘it is necessary that’. The grammatical subject *il* was dropped in (8) while the complementizer *que* was dropped in (9). As a rule, the deontic nature of *falloir* triggers a subjunctive verb form in the embedded clause, as in (8). However, as shown in (9), this rule is not sacrosanct. Poplack et al. (2013) looked into a number of factors influencing the choice of mood after *falloir* and other matrix verbs. Their study showed that the rate of subjunctives decreases with increasing

distance between the controller verb and the target verb. Unfortunately, distance was measured in a rather coarse-grained fashion (e.g. in terms of the presence or absence of parenthetical material). This makes it difficult to gain a precise idea of the impact of distance on mood choice. Nonetheless, we may provisionally interpret the French data to mean that the choice of the subjunctive depends on the availability of a nearby trigger whereas the choice of the indicative is insensitive to the distance of the trigger.

Variation (such as that between the indicative and the subjunctive) has been customarily understood to provide evidence of a set of factors which may generate alternative outputs (see Hilpert 2008 for many others). One factor may facilitate output A while another may facilitate output B. Which of the alternatives A and B is chosen depends on the relative strength of these influences. The strength of a factor varies as a function of the relationship it contracts with other factors at the moment of production planning. It may also vary with historical time (e.g. Gries & Hilpert 2010) and from dialect to dialect (e.g. Szmrecsanyi et al. 2016).

The vast variationist literature has conceived of factors in an essentially static manner. In particular, the temporal dimension has been largely ignored.² However, once these factors are assumed to have some psychological reality, we cannot afford to leave the time factor out of consideration. In language production, units unfold in real time. They have a characteristic time course of activation, with a beginning and an end as well as a particular strength at any one moment in time. The 'life span' of a unit will be dubbed its RANGE. Once alternative options are 'contemporaries', the possibility of competition arises. As in the case of the strength of an individual unit, competition between alternatives varies from one moment to another.

It is the aim of the present article to explore the notion of range in some detail and thereby enrich current modelling of linguistic variation by the temporal dimension. Distance effects are ideally suited to studying the range of linguistic units. As discussed above, they show how different decisions may be taken at different moments in time and thereby provide insight into the time course of activation of linguistic units.

In this paper, the concept of range will be applied to factors which play a part in grammatical variation. The basic idea is that different units may have different ranges. An attempt will be made to identify the principles which influence the range of a given unit and to develop a unified model of how distance effects come about. The grammatical issue on which the present work is based is the English

[2] In psycholinguistics, timing is of course a familiar concept. Priming effects have always been understood to be sensitive to time, i.e. subject to decay (see e.g. Swinney 1979, Branigan, Pickering & Cleland 1999, Jaeger & Snider 2008). Priming has not only been shown to affect participants' reaction times in psycholinguistic experimentation but also to influence the choice of structural options in text corpora (e.g. Gries 2005, Szmrecsanyi 2006, Reitter, Moore & Keller 2006). In these works, priming is one variable among others. By contrast, our approach is not concerned with priming. As we explain in what follows, we are interested in the availability of linguistic units as a function of time.

mandative subjunctive. As will be explained in the next section, the subjunctive is an especially rewarding area to probe because it offers a ternary rather than a binary choice, the latter of which is characteristic of most other variationist phenomena.

2. THE INTEREST OF THE MANDATIVE SUBJUNCTIVE

As a conceptual category, the mandative subjunctive in Modern English can be expressed in the three ways illustrated in (10) from Greenbaum & Quirk's (1990: 44) standard reference work.

- (10) (a) The employees demanded that he resign.
 (b) The employees demanded that he should resign.
 (c) The employees demanded that he resigns.

The conceptual category is coded by the subjunctive in (10a), a modal auxiliary in (10b) and the indicative in (10c). The fact that the three options are compatible with the verb *to demand* demonstrates that the form of the verb in the subordinate clauses is not uniquely determined by a particular trigger. So there is room for grammatical variation. By common consent, the decision among these variants is influenced by region, mode (spoken vs. written language) and (historical) time (e.g. Övergaard 1995, Serpollet 2001, Leech et al. 2009). We remain agnostic as to the role of semantics in this game – apart from the obvious requirement that the trigger of the mandative subjunctive have a deontic meaning. While it has been claimed that the different options are not semantically equivalent (e.g. Övergaard 1995: 11; Hoffmann 1997), the meaning differences appear to be rather minor. A connection between mandatory force and mood choice has not as yet been established.³ In any event, as long as the status of the three forms in (10) as variants is not in dispute, the role of semantic contrast is largely inconsequential for our investigation (see Hinrichs, Szendrői & Bohmann 2015: 814). This follows from the lack of a link between meaning and distance. Potential semantic or pragmatic differences cannot naturally predict potential differences in range.

The three options in (10) allow us to address two major issues. The first is the structural level at which the mandative subjunctive is coded. While the indicative is expressed morphologically (by means of a suffix, which can also be zero), the modal auxiliary is a free-standing, syntactically manipulable unit and therefore qualifies as a syntactic option. We may thus compare the range of the syntactic level to that of the morphological level within a structural model which recognizes

[3] However, there is a link between mandatory force and the choice of the modal auxiliary called modal harmony in Huddleston (2002: 997). Verbs with weak mandatory force (e.g. *to suggest*) select weak modals like *might*, for example, more often than do verbs with strong mandatory force (e.g. *to insist*) (Vlasova 2010). See also Bybee, Perkins & Pagliuca (1994: 214) who view the emergence of *should* as satisfying a harmony constraint between the deontic nature of the modal and that of the matrix verb.

different levels but also provides for a good deal of interactivity between them. Given that the syntactic level is superordinate to the morphological level in the structural hierarchy, this comparison provides an opportunity of examining the impact of hierarchical position on range. It might be that higher levels have a wider range than lower levels because high-level planning has to oversee larger domains than low-level planning.

The second issue emerges from the similarity and contrast between the indicative and the subjunctive. With a suffix or the absence thereof as their exponents, both mood values originate at the morphological level and hence allow us to study within-level effects. At the same time, the fact that the indicative is sometimes realized by a suffix while the subjunctive is realized by zero, grants insight into the contrast between explicit and implicit coding. This may not be unlike the contrast between the prepositional verb *to stop OBJ from Xing* and the prepositionless verb *to stop OBJ Xing* discussed in the opening section. If the syntactic data from the *stop*-construction generalize to the morphological level, we would expect a shorter range for implicit than for explicit coding.

To conclude, the English subjunctive gives us the welcome opportunity of studying the following three issues: by contrasting the modal strategy with the other two options, we can test for possible differences in range between the syntactic and the morphological level. In addition, by contrasting the subjunctive with the indicative, we can test for possible differences in range between an explicit and an implicit coding strategy. Finally, by comparing these two analyses, we can test for possible differences in the size of between-level and within-level effects. However, before we can present the empirical analysis, a variety of methodological points have to be taken care of.

3. METHOD

Most usually, the subjunctive depends on an explicit trigger (but see Hundt 1998a for some qualification). Our first task was therefore to identify subjunctive triggers. An attempt was made to set up a (more or less) exhaustive list.⁴ Our starting point was Crawford's (2009) Appendix A, which includes 47 verbal, 38 nominal and 23 adjectival triggers. Three types of change were made to this list, one substitutive, one additive and one subtractive. Crawford's list contains the verb *to intimidate* which we are unsure about. It might be an inadvertent substitution for *to intimate*. Be that as it may, we tested for subjunctive uses of *to intimidate* but did not find any. So we decided to replace *to intimidate* with *to intimate*. We deleted all items from Crawford's list for which our own analysis

[4] Of course, we do not claim to come up with a complete list. It is highly likely that such a list can never be complete because speakers and writers may occasionally use a trigger deontically which is most usually used factually. While we cannot rule out the possibility that we may have overlooked a few triggers, we are confident that the inclusion of these uncommon cases would not significantly alter our patterns of results.

(see below for details) did not yield any deontic uses at all. This was true for one adjective (i.e. *convenient*), five nouns (e.g. *implication*) and four verbs (e.g. *to insure*). At the same time, in the course of our research, we came across as many as 21 adjectives (e.g. *adamant* and *critical*) and one noun (i.e. *urge*), which were not on Crawford's list but had the potential to trigger a subjunctive. With these changes, we ended up with a total of 120 subjunctive triggers, which divide into 43 verbs, 34 nouns and 43 adjectives. The complete list is provided in the appendix.

There are four signposts to signal a subjunctive in English. Apart from the absence of the third person singular present tense inflection, as in (10a), there is the subjunctive form *be* from the infinitive *to be*, the lack of *do*-support in negated clauses and the non-occurrence of backshifting of tense. The latter three options are exemplified in (11)–(13) from our corpus. For easy identification, trigger and target will henceforth be put in bold.

- (11) It is **recommended** that the walls **be** decorated with pictures of cartoon characters.
- (12) Curiously, this passage occurs immediately after a **demand** that women **not adorn** themselves with jewels and braided hair!
- (13) But at that time both leaders felt it was very **important** that they **speak** to one another.

The exceptional status of *to be* facilitates the identification of the subjunctive because the subjunctive form *be* differs from the indicative forms in the entire person paradigm, as exemplified by the third person plural subject in (11). The fact that the subjunctive in negated clauses does not require the prop word *do* (as in (12)) is another indication of a subjunctive form. The lack of backshifting of tense can also be regarded as a clue to the subjunctive. In (13), the narrative is in the past tense. We would therefore expect the verb form in the subordinate clause to be *spoke* if it were indicative. The fact that this does not occur implies that *speak* is a subjunctive. However, this logic can only be reliably applied when the verb in the matrix clause is in the past or perfect tenses. When the target verb is preceded by a non-finite trigger, this criterion is obviously of no avail.

In view of the almost complete lack of formal marking of the subjunctive, verb forms ambiguous between the indicative and the subjunctive abound. One such case is (14).

- (14) Many medical boards also **require** that specialists **get** recertified every 7 to 10 years.

As all non-third-person-singular forms of verbs other than *to be* do not formally distinguish between indicative and subjunctive mood, the finite form *get* in (14) could be either one or the other.

Owing to polysemy, one and the same trigger can have a deontic or a factual reading. In fact, almost all triggers on our list which are not too uncommon allow both uses, though to widely varying degrees. The two uses are exemplified in (15) on the basis of the trigger *to suggest*.

- (15) (a) All the **evidence** suggests that he **has** stolen the money.
 (b) He **suggested** that she **eliminate** all possible dietary triggers.

The epistemic meaning of *to suggest* in (15a) contrasts with its deontic meaning in (15b). In the vast majority of cases, the decision as to whether an individual token of our trigger types is used deontically or factually was a straightforward one. In making this decision, we could rely on a number of clues including the meaning of the trigger and the use of the subjunctive or a modal auxiliary. In case of indicative use, the context usually clarified whether a deontic meaning was intended. More particularly, we relied on certain clues such as the temporal relationship between the situation described by the matrix clause and that described by the subordinate clause. A deontic function can only refer to situations which follow the situation referred to. Take the following example:

- (16) And that's not the case with any modern nuclear reactor. So, it is **crucial** that this **was** a 1971 machine.

The machine in question has always been from 1971, which predates the particular situation described here (the 2011 Fukushima nuclear disaster). This establishes the situation described by the subordinate clause as a fact, which does not lend itself to a deontic interpretation. That is, the reactor cannot be mandated to be from any year other than 1971, and the speaker obviously did not intend to do so. More generally, then, a tell-tale sign of a factual use is when the situation described by the subordinate clause chronologically precedes or coincides with that of the matrix clause.

Another clue is provided by the nature of the subject NP. Consider (17).

- (17) Rationality and human experience **demanded** that he **couldn't** exist.

As abstract concepts, neither rationality nor experience can actually demand that someone do something. Hence, the meaning of *demand* in (17) cannot be deontic; in fact, it is very similar to the non-deontic meaning of *suggest* in (15a), where the subject is also an abstract noun. This automatically forces a more factual interpretation in (17): what other people are capable of cannot be dictated.

In the few cases where we were in doubt even after applying these criteria, our practice was to err on the side of caution and classify them as factual.

We move on to a description of our database. The (online version of the) Corpus of Contemporary American English (COCA) from 1990 to 2012 was selected mainly for two reasons. In addition to its sheer size, we preferred to examine American rather than British English because the mandative subjunctive occurs

more commonly in the former than the latter variety (e.g. Erdmann 1981, Hundt 1998a, Crawford 2009).

COCA was searched for all the triggers on our list. The search term was ‘trigger (in all its inflectional variants) + *that*’. Hence, contact clauses (e.g. *At the peak of Ellen Pyle’s flourishing career in the 1920s and 30s, the Post editor requested she write her life story for publication in the magazine*) and cases where triggers and complementizers were non-adjacent (e.g. *But I mean, it’s really important psychologically that we go up tomorrow because, otherwise, this is very painful*) were not retrieved. The exclusion of contact clauses and non-adjacent triggers and complementizers was motivated by both practical and empirical reasons. If the search term had been restricted to the bare trigger or had been wide enough to include intervening material, the amount of manual editing would have been substantial. Moreover, contact clauses constitute a minority in mandative constructions (e.g. Johansson & Norheim 1988, Hundt 1998a, Waller 2017). The same is true of cases with non-adjacent triggers and complementizers.

For many triggers, the COCA search yielded fewer than 1,000 hits. All tokens of these verbs were subjected to scrutiny. For practical reasons, triggers for which COCA produced more than 1,000 hits were limited to a sample of 1,000 hits. COCA generates such random samples, including tokens from all years and subgenres.

The computer-assisted search turned up a number of hits where the matrix verb governed more than one embedded verb, as documented in (18).

- (18) It is also **essential** that they **receive** equal pay for equal services, **be** provided adequate working space, **benefit** from due procedures, and **be** allowed equal opportunity for professional development.

As many as four verbs depend on the trigger *essential*. All these subordinate verbs were treated as separate hits because the distance between trigger and target differs from case to case. The decision for counting the individual targets separately was additionally motivated by the fact that the subordinate verbs did not necessarily have the same mood value (even though they depend on the same trigger). This suggests a certain autonomy of the individual clauses with respect to mood choice. The net effect of this procedure is that some triggers have slightly more than 1,000 hits in our corpus.

Hardly any editing of the corpus was performed. Only when the exact same string of words from the same source and the same year accidentally appeared as two or more different hits did we eliminate the duplicates. Some hits were marred by minor errors such as when *she* appeared as *sbe*. As the context almost always made clear which word was meant, these errors were rectified and the hit correctly coded. The very few cases of mistranscription which produced incomprehensible utterances were discarded.

Next, we turn to our data coding scheme.⁵ Each token was coded for twelve variables – some straightforward, some necessitating some special consideration. The rest of this section introduces and discusses them in turn.

In addition to the trigger itself, its deonticity, its textual frequency and its word class was noted. Deonticity is a binary variable (yes/no). The word-class variable has three values: verb, noun or adjective. The textual frequency of the trigger was determined on the basis of its occurrences in COCA. Trigger frequency was operationalized by calculating the token frequency of the trigger alone as well as the string frequency of the trigger and the following complementizer.

Kastronic & Poplack (2014) caution that much of what we know about the mandative subjunctive (and its history) derives from the analysis of written language and that it is anything but obvious that these results generalize to speech. It was deemed necessary therefore to distinguish between spoken and written utterances. COCA includes both spoken and written language and thus allows us to test whether this factor plays a role in the subjunctive game and, more specifically, whether possible distance effects can be observed in both modes. However, it should be acknowledged that the distinction between spoken and written language in general and in COCA in particular is not at all clear-cut. The spoken sample in COCA is based on media speech (e.g. talk shows on TV) and hence does not reflect completely natural and spontaneous conversation.

The next variable looks into the mood of the subordinate verb. We observed five values in the data: modal, subjunctive, indicative, ambiguous and imperative. A separate column was reserved for the modal auxiliaries. Technically speaking, this column is nested in the modal value of the Mood column. As many as 13 modals or functionally similar verbs such as *to have to* and *to be to* were found to occur in subordinate clauses.

The critical variable in this study is the distance between trigger and target. We decided to measure distance linearly⁶ in terms of number of words as well as number of morphemes. In the ideal case, this double strategy would allow us to argue that possible distance effects are independent of the particular measure being chosen.

[5] The raw data underlying this study may be obtained upon request.

[6] Our decision in favour of distance as a linear principle aligns our work with much previous research (e.g. Hawkins 1994, Gries 2003, Köpcke et al. 2009). Of course, distance can also be conceived in hierarchical (or temporal) fashion. It is now fairly certain that neither the linear nor the hierarchical notion can be reduced to the other (e.g. Bock & Cutting 1992, Vigliocco & Nicol 1998, Wasow & Arnold 2003, Berlage 2014). Despite the irreducibility of the linear and the hierarchical principles, there is wide agreement that the two operationalizations of distance measure highly similar things (e.g. Wasow 1997, Ferreira 1991 and in particular Szmrecsanyi 2004). Both measures are similarly successful in accounting for the same experimental data (see Santiago et al. 2000 vs. Roelofs 2002 vs. Santiago, MacKay & Palma 2002). This is not really surprising. The greater the linear distance between two given units, the greater the complexity of the structural representation. The high degree of co-variation between the two measures justifies focusing on only one of them.

Obviously, the count of words and morphemes requires working definitions of what counts as a word or a morpheme. To make the separate counts more meaningful, our overall strategy was to maximize the difference between the two types of unit. That is, units which were difficult to classify were treated as single words in the word count but as multiple morphemes in the morpheme count. Accordingly, compounds (whether hyphenated, written solid or written separately), complex proper nouns (e.g. *General Motors*) and contracted forms, in particular negated forms such as *don't* and *wasn't*, were counted as one word in the word count. However, first and last names were counted as one word each.

Every morpheme that was relatively clearly identifiable was included in the morpheme count. Minor morphophonological variation (as in *heal* – *health*) was ignored. That is, *health* was counted as bimorphemic, as were irregular past tense forms such as *sang* and *felt*. However, we did not take the morpheme maximization strategy to its extreme. Derived words with no semantic relation to their base (e.g. *department* – *to depart*), derived words with a considerably higher frequency than their bases (e.g. *dictionary* – *diction*) as well as seemingly derived words with virtually no existent base (e.g. *environment*) were classified as monomorphemic. This is to do justice to their holistic representation in the mental lexicon (e.g. Bybee 1985, Marslen-Wilson et al. 1994, Hay 2001). Words containing cranberry morphemes (e.g. *aftermath*) and suppletive forms were treated likewise. A surface approach was adopted according to which only 'visible' morphemes were taken into consideration. An example such as *Americans'* in *Americans' health insurance* was counted as a three- rather than a four-morphemic word. Words ending in *-ical* require special mention. This ending was usually categorized as bimorphemic because there is a lexical contrast between, let us say, *economic* and *economical*. However, this decision did not seem appropriate for words like *radical*, for example, where the holistic touch predominates.

It turned out on later analysis that the word and the morpheme count could not be included as independent factors in the same statistical model because they are correlated to an extremely high degree ($r_{\text{pearson}} > .9$) (see also Yaruss 1999). We therefore decided to design two separate models, one including distance in words and the other distance in morphemes, and to examine which of the two operationalizations resulted in a better model fit. Not surprisingly, the performance of the two models was almost identical. The final model featured words as the relevant measure of distance because it yielded a minimally (non-significantly) better model fit. Thus, distance in morphemes is not given any attention in our presentation of the statistical results.

Our general strategy in measuring distance was to count the material intervening between target and trigger. In the case of the modal option, the modal auxiliary was regarded as the target and accordingly ignored. A particular challenge was presented by negation. The problem was whether the negation marker *not*, which is supported by *to do* in an indicative construction but not so in a subjunctive construction, should be counted in or out. Contrast (19) with (20).

- (19) Hlonipha **requires** that she not **look** her father-in-law in the eyes.
- (20) They **require** that I **do** not lessen or make light of our history simply because . . .

We decided against including *not* in our distance counts because the occurrence of the negation marker is not independent of mood choice. Viewed from a linear perspective, at the moment that *not* appears in (19), the decision in favour of the subjunctive has already been taken. As noted above, the unsupported *not* is a special property of the subjunctive, so neither an indicative nor a modal construction may follow. Therefore, the negation marker *not* was not included in the distance counts. In fact, it was completely ignored because it invariably follows the modal auxiliaries (e.g. *should not*) and finite verbs (see (20)).

The distance between trigger and target was not only coded in terms of number of words and morphemes but also in terms of number of additional clauses, both finite and non-finite. The ordinary case was a trigger followed by the complementizer *that*, the subject of the embedded clause and the target. However, syntactically more complex constructions, in particular relative, conditional and complement clauses, occasionally found their way into the data, as exemplified in (21).

- (21) It is also **necessary** that he who makes the atonement **should** be the very being who has sinned.

While we originally intended to examine the syntactically more complex cases separately, the later analysis yielded an insufficient number of relevant items for statistical testing. Therefore, this dimension was not considered further.

The last two variables that were taken into account are the nature of the target verb, in particular whether or not it is a form of *to be*, and the presence or absence of the negation particle *not* in the subordinate clause. While we do not put forward any hypotheses as to a possible influence these variables might exert on mood choice, we decided to include them because of the special role they play in the subjunctive construction (see above).

4. DATA ANALYSIS

The selection procedure depicted in the previous section yielded a total of 49,154 data points, which divide into 17,626 deontic and 31,528 factual uses (35.9–64.1%). The factual uses were excluded from further analysis. The imperatives were so extremely infrequent ($N = 12$) that they were also discarded. [Table 1](#) presents a descriptive survey of the remaining 17,614 data points. For the sake of completeness, it reports both the word and the morpheme count.

Prior to looking at distance effects, it is appropriate to point out that [Table 1](#) confirms the frequency distribution of the various mood values reported in the relevant literature (e.g. Johansson & Norheim 1988, Hundt 1998b, Schneider 2005). In American English, subjunctives are clearly the majority choice (56.0%)

Mood choice	Average distance in. . .	
	words	morphemes
Modal (N = 2,406)	5.65	7.70
Indicative (N = 565)	4.44	6.37
Subjunctive (N = 9,861)	4.30	5.78
Ambiguous (N = 4,782)	4.19	5.86

Table 1
Mood choice as a function of distance (N = 17,614).

even if ambiguous cases are counted in. If ambiguous cases are ignored, the rate of subjunctives rises to 76.8%. Modals are much less frequent, accounting for 13.7% (including ambiguous cases) or 18.8% (excluding ambiguous cases). The indicative is an unlikely choice at 3.2% (ambiguous cases included) or 4.4% (ambiguous cases excluded). The rather high rate of ambiguous cases is an immediate spin-off function of the impoverished exponence of the subjunctive (see Section 3).

Table 1 shows that the average distance between trigger and target is highest for modals, lower for indicatives and even lower for subjunctives. This difference is much larger for the contrast between modals and non-modals ($5.65 - 4.28^7 = 1.37$ words; $7.7 - 5.83 = 1.87$ morphemes) than for the contrast between indicatives and subjunctives ($4.44 - 4.3 = 0.14$ words; $6.37 - 5.78 = 0.59$ morphemes). As can be seen, the word and the morpheme counts yield very similar rankings.

The ambiguous cases are very close to the subjunctive. In fact, the difference in distance between ambiguous cases and subjunctives is smaller than that between ambiguous cases and indicatives. Coupled with the observation that indicatives are a rare choice in our data, this strongly suggests that ambiguous cases are more likely to be covert subjunctives than covert indicatives.

We move on to a statistical treatment of the data. Given the three mood values, two binary comparisons were performed: the between-level comparison contrasts the syntactic with the morphological cases and the within-level comparison contrasts the subjunctive with the indicative mood. The within-level comparison is subordinated to the between-level comparison since the former takes a more detailed look at the (unambiguous) morphological options, which the latter treats as an undifferentiated set.

Within each comparison we carried out two analyses. First, in order to get an overall idea of the distance effect, we calculated the ratios of the competing mood choices by distance between trigger and target and analyzed the correlation between the two by means of a regression analysis. In a second step we built

[7] The values of 4.28 and 5.83 are derived from averaging across indicatives, subjunctives and ambiguous cases.

mixed-effects logistic regression models, which allow us to test the effect of distance on mood choice, while controlling for the variables we introduced in the previous section.

4.1 *Between-level comparison: Syntactic vs. morphological*

For the between-level analysis, modal constructions were classified as syntactic but all others as morphological. The ambiguous cases were assigned to the morphological set because they are ambiguous between indicative and subjunctive, both of which belong to the morphological level.

4.1.1 *Ratio analysis*

What effect does distance have on the choice between the syntactic and the morphological variants? In an attempt to answer this question, we calculated the ratio of the syntactic vs. the morphological option by distance (measured by the number of words). We excluded distances between trigger and target in excess of 15 words. This region yielded a low number of data points, which made it difficult to reliably estimate the syntactic/morphological ratio, that is, to generate a robust statistical model. This decision resulted in a loss of 574 data points (3.3% of the data). The diagram in [Figure 1](#) plots the syntactic/morphological

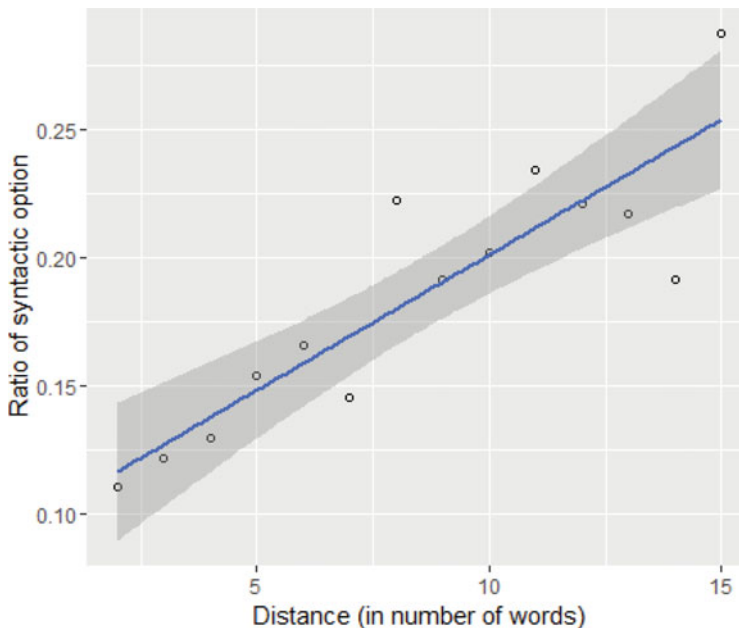


Figure 1
Syntactic/morphological ratio by distance between trigger and target.

ratio on the y-axis versus the distance between trigger and target on the x-axis. The regression line in this figure shows quite clearly that the ratio of the syntactic variant increases with increasing distance. A regression model predicting the ratio of the syntactic variant by distance in number of words yields the following output (see Table 2).

Variable	Coefficient estimate	Standard error	<i>t</i>	<i>p</i>
Intercept	0.0955	0.0153	6.260	< .001
Distance (in number of words)	0.0106	0.0016	6.516	< .001

Table 2

Output of linear regression model predicting the ratio of the syntactic option (N = 17,040).

The model output indicates a highly significant effect of distance on mood choice. The coefficient of 0.0106 means that the likelihood of the syntactic option increases by about 1% with each increment of one word. The adjusted *R*-squared value of the model is 0.76, which indicates a very high correlation between ratio and distance ($r_{\text{pearson}} = .88$). Thus, the statistical model predicts the empirical patterns quite well.

4.1.2 A mixed-effects analysis

Two limitations of the above model are its monofactorial nature and the fact that it is based on mean values but does not capture decisions at the level of individual tokens. Since it is unlikely that distance is the only explanatory factor in this game, it is necessary to perform a second statistical analysis which examines whether the choice between the syntactic and the morphological option is under the sway of distance, even when possibly confounding variables are taken into account. To this end, a mixed-effects logistic regression model fitted to the choice between the syntactic or the morphological variant was calculated. This model is based on the same sample as the ratio model, with distances larger than 15 words excluded.

In addition to Distance, our model includes as fixed effects the following five variables: Trigger Frequency, Part-of-Speech (POS), Mode, Negation and Be, all of which were discussed in Section 3 above. POS is a ternary variable (noun, verb, adjective) while the others are binary: Mode (spoken vs. written), Negation (yes vs. no) and Be (yes vs. no). The predictors Distance (in number of words) and Trigger Frequency were log-transformed before being entered into the model because their distribution was found to be more normal in logarithmic space.

The random effects structure of the model includes random intercepts for trigger and corpus file, which we use as a proxy to control for speaker/writer-specific effects. We furthermore added random slopes by trigger and by corpus file for the critical variable Distance.

Throughout the model fitting procedure, we kept the random-effects structure maximal, following a design-based approach (see Barr et al. 2013). Models were built using the *glmer* function⁸ of the *lme4* package (Bates et al. 2015) in R (R Development Core Team 2014). With regard to the fixed effects, we first built a maximal model featuring all aforementioned predictors as main effects as well as all pairwise interactions of Distance with the other predictors. We subsequently removed interactions and fixed effects in a stepwise fashion, beginning with the term with the highest *p*-value.

The model fitting process revealed an interaction effect between the predictors Distance and Negation, which, however, led to nonconvergence of the corresponding models. An examination of the data revealed that this was due to a shortage of data points with negated subordinate clauses, for which the distance effect could thus not be properly estimated. In order to steer clear of potentially damaging effects of Negation on our model, we decided to eliminate all cases with negated subordinate clauses. This resulted in a loss of 938 data points (5.5% of the data).

Model fitting on the reduced data set showed none of the other two-way interactions with Distance to be statistically significant. Because the predictors Be and Trigger Frequency were not significant as main effects, they were discarded. Table 3 provides the complete model output of the final model and should be read as follows: positive coefficients indicate a heightened probability of choosing the syntactic option while negative coefficients indicate a heightened probability of choosing the morphological option. For the ternary variable POS the variable value 'noun' served as a baseline.

The resultant model is a minimal adequate model with all fixed-effect predictors contributing significantly to model fit. The major result is a highly significant effect of Distance. The probability of choosing the syntactic option increases with increasing distance between trigger and target. This effect is robust enough to show up in the presence of other variables which also exert a significant influence on mood choice.

For the variable POS, we observe that while there is no significant difference between nouns and verbs, nouns are more likely than adjectives to choose the syntactic variant. Mode also impacts on Mood choice: the syntactic option is significantly more likely to occur in speech than in writing.

As regards the random effects, the random intercept for trigger accounts for a considerably larger share of variance than the random intercept for filename. This shows that the choice between the syntactic and the morphological option

[8] We ran the *glmer* function employing the *bobyqa* optimizer function.

Random effects	Variance	Standard deviation		
Corpus file (intercept), 1,000 groups	0.01501	0.1225		
Corpus file (distance)	0.03801	0.1950		
Trigger word (intercept), 120 groups	3.64677	1.9097		
Trigger word (distance)	0.12575	0.3546		

Fixed-effect predictors	Coefficient estimate	Standard error	<i>z</i>	<i>p</i>
(Intercept)	-1.29043	0.34688	-3.720	.0002
POS = Verb	-0.36459	0.42358	-0.861	n.s.
POS = Adjective	-1.18489	0.43733	-2.709	.0067
Mode = written	-0.22064	0.07681	-2.873	.0041
log-transformed Distance	0.43274	0.07244	5.973	.0000

Table 3

Output of mixed-effects model predicting the odds of choosing the syntactic or the morphological option (N = 16,102).

is driven to a large extent by the individual trigger, but not so much by idiolectal variation.

The bottom line of the mixed-effects analysis is that it fully corroborates the ratio analysis, with the effect of Distance emerging in both analyses. The fact that none of the two-way interactions of Distance and the other predictors proved significant demonstrates the autonomy of the distance effect.⁹

4.1.3 Multiple targets

It was mentioned in [Section 3](#) that our corpus contains sentences in which more than one clause is subordinated to the matrix clause (see (18) above). These cases allow us to address two additional questions. Is mood choice consistent across the different target verbs? In other words, does the mood of the first subordinate verb determine the mood of all following verbs? If not, is the switch from one mood value to another symmetrical, i.e. are switches from morphological to syntactic as likely as the reverse direction? Our previous results lead to a very clear prediction. Since the distance between trigger and target increases considerably with each additional subordinate clause, a switch from morphological to syntactic should be significantly more frequent than a switch from syntactic to morphological.

[9] We also built a statistical model which included Genre as a predictor variable (with the following values in COCA: Spoken, Academic, Newspaper, Fiction, Popular Magazines). None of the text types interacted with Distance. This is true of both the between-level comparison and the within-level comparison to be reported in [Section 4.2](#).

We have found 1,396 triggers with more than one target in our data. Of these, 57 (= 4.1%) involve a switch from one mood value to another. This rather low percentage indicates that switches are a dispreferred option. The mood of the first target is a highly reliable predictor of the mood of the second (and following) target verb. Two exceptions to this pattern are given in (22) and (23).

- (22) It is therefore manifestly **desirable** that the belief either **be** proved false, or **should** be confirmed . . .
- (23) Japan's leaders were **determined** that no such regretful wakes **would** be held for them and that they not **be** denied their 'proper place' in Asia.

While (22) exemplifies a change from the morphological to the syntactic option, (23) illustrates a change in the opposite direction from the first to the second target verb. Of the 1,396 sentences with at least two subordinate clauses, 47 involve a switch from morphological to syntactic but only 10 a shift from syntactic to morphological. This difference is highly significant (binomial test, $p < .0001$). This result provides further support for the claim that the likelihood of the syntactic option rises with increasing distance between trigger and target.

4.2 *Within-level comparison: Indicative vs. subjunctive*

We proceed to a more detailed examination of the morphological level. The elimination of the ambiguous cases leaves us with the choice between indicative and subjunctive mood. A closer look at the two options affords us an opportunity of testing within-level effects. Does the indicative have the same range as the subjunctive effect? If Poplack's (1992) data from French (discussed in the opening section) carry over to English, the answer might be in the negative. In that case, a longer range would be predicted for the indicative than the subjunctive. As in the between-level comparison, both a ratio and a mixed-effects analysis were carried out.

4.2.1 *Ratio analysis*

For the analysis of the effect of distance on the choice between indicative and subjunctive, items with distances of more than 15 words were discarded for the same reasons as stated above. This resulted in a loss of 311 data points (3.1% of the data) and left us with 547 indicatives and 9,568 subjunctives.

The diagram in [Figure 2](#) plots the indicative/subjunctive ratio on the y-axis and the distance between trigger and target in number of words on the x-axis. [Figure 2](#) suggests an effect of Distance on the choice between indicative and subjunctive mood. As the regression line indicates, the indicative/subjunctive ratio increases with increasing distance between trigger and target. The output of the corresponding regression model is reported in [Table 4](#).

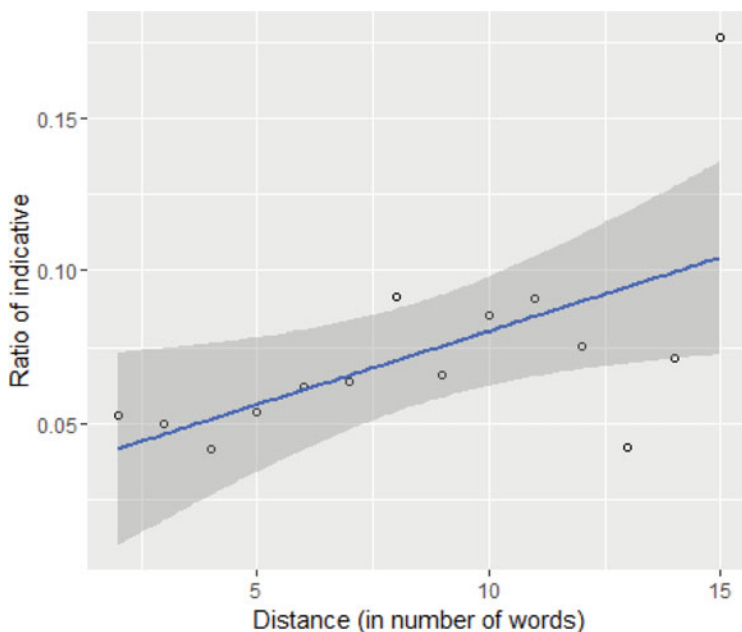


Figure 2
Indicative/subjunctive ratio by distance between trigger and target.

Variable	Coefficient estimate	Standard error	<i>t</i>	<i>p</i>
Intercept	0.0323	0.017795	1.816	n.s.
Distance (in number of words)	0.0048	0.001892	2.538	< .05

Table 4
Output of linear regression model predicting the indicative vs. subjunctive ratio (N = 10,115).

Table 4 shows that the regression line in Figure 2 reflects a significant effect of distance on the indicative/subjunctive ratio. The coefficient of 0.0048 indicates that the probability of the indicative increases by approximately 0.5% per word increment in distance. The adjusted R-squared value is .295, the correlation coefficient is $r_{\text{pearson}} = .59$, indicating a moderate correlation between distance and mood choice.

4.2.2 A mixed-effects analysis

As before, we calculated a mixed-effects model with an eye to investigating whether Distance preserved its significant influence on mood choice in the face

of competition from other variables. The random-effects structure of the model was the same as above, with random intercepts for trigger and filename as well as random slopes of Distance by trigger and filename. As the models with random slopes failed to converge, random slopes had to be left out of account.

The same fixed-effects predictors were subjected to testing as in the previous set of analyses. We tested Distance, Trigger Frequency, POS, Be and Mode as main effects as well as all pairwise interactions of Distance and the other factors. For the same reason as given in Section 4.1.2, the items with negated subordinate clauses were removed. This led to an exclusion of 726 data points (7.2% of the data).

The model output is presented in Table 5. Positive coefficients indicate an elevated likelihood of selecting the indicative while negative coefficients indicate an elevated likelihood of selecting the subjunctive.

Random effects	Variance	Standard deviation		
Corpus file (intercept), 1000 groups	0.3585	0.5987		
Trigger word (intercept), 107 groups	1.4980	1.2239		

Fixed-effect predictors	Coefficient estimate	Standard error	<i>z</i>	<i>p</i>
Intercept	-4.7908	0.4318	-11.095	.0000
POS = Verb	-0.7448	0.4684	-1.590	.1118
POS = Adjective	1.7409	0.4440	3.921	.0001
Mode = Written	-0.4848	0.1417	-3.421	.0006
BE	1.1229	0.2812	3.993	.0001
log-transformed Distance	0.6554	0.1593	4.115	.0000
BE:Distance	-0.5589	0.2050	-2.726	.0064

Table 5

Output of mixed-effects model predicting the odds of choosing the indicative or the subjunctive (N = 9,389).

The most important result to emerge from Table 5 is the highly significant main effect of Distance, which seconds the output of the ratio model. The probability of choosing the indicative increases with increasing distance between trigger and target. With the exception of Trigger Frequency, all other main effects also turn out to be significant. While adjectives are significantly more likely to trigger the indicative than are nouns, the difference between verbs and nouns is not statistically significant. When the target verb is a form of *to be*, the likelihood of the indicative is increased. The coefficient for Mode reveals that the subjunctive is more typical of the written than the spoken language (in line with Peters 2009).

Table 5 furthermore shows that Distance is involved in a statistically significant interaction with Be. The net effect is that Distance plays less of a role in mood choice when the verb in the subordinate clause is a form of *to be*.

Concluding, Distance emerges as a reliable predictor in the choice between indicative and subjunctive mood. The indicative tolerates larger distances better than the subjunctive does. This conclusion is consonant with Poplack's results for French. What makes this consonance especially remarkable is that the English and the French subjunctive differ in form, function and frequency of use.

4.3 Comparing between-level and within-level effects

A comparison of the between-level and within-level analyses reveals that Distance plays a significant role in both. It remains for us to probe into a possible interaction of type of analysis and size of the distance effect. Going by the coefficients of the regression models based on the ratios of the two options (see Tables 2 and 4), the between-level effect is twice as strong as the within-level effect. We tested for significance of this difference by calculating a z -test as suggested by Clogg, Petkova & Haritou (1995: 1276), which yields a statistically significant result ($z = 2.313, p < .05$). We conclude that the distance effect is stronger in the between-level scenario compared to the within-level analysis.

5. THEORETICAL DISCUSSION

The analysis of the English subjunctive has shown that the likelihood of choosing a particular mood value varies with the distance between trigger and target verb. The larger the distance, the more likely the syntactic option in the between-level analysis and also the more likely the indicative option in the within-level analysis. Implicationally, the syntactic level has a wider range than the morphological level, and the indicative has a wider range than the subjunctive. Owing to this difference in range, syntax can support the production of a modal auxiliary at greater distances more easily than morphology can support the production of the indicative or the subjunctive; hence there is an increasing probability of choosing modals at larger distances. The difference in range is without effect when the distance between trigger and target is small. At short lags, syntax and morphology can easily perform the same task even though their strengths need not be identical.¹⁰ The same reasoning holds for the within-level analysis. The wider range of the indicative is conducive to an increasing likelihood of selecting

[10] This overlap is as predicted by a parallel interactive processing model. Note also that this job sharing is a probable reason for the typological fact that one and the same problem may be solved syntactically in one language but morphologically in another. As one example for many others, take case languages such as Hungarian and Finnish which express the same information morphologically which is coded phrasally in other languages. See Haspelmath (2011) who argues against the possibility of drawing a clear dividing line between morphology and syntax in language typology.

the indicative at larger distances. At short lags, however, the indicative and the subjunctive can in principle do the same job.

Rather simplistically, we may divide the time course of activation of linguistic units into three parts: a rise rate in the initial stages, an average activation level in the intermediate stages (which is dynamic rather than static) and a decay rate in the final stages. At this point of our paper, we have nothing to say about the beginning of the activation process (but see next section). Our data speak more directly to the middle part. The fact that the morphological option is generally preferred in our data suggests that morphology is stronger than syntax, in particular for the more common distances between trigger and target. By the same token, the fact that the subjunctive is generally more frequently chosen than the indicative lends support to the hypothesis that the subjunctive attains a higher activation level than the indicative most of the time.

The notion of range speaks most directly to the final part of the activation process. Varying ranges can be translated most readily into varying decay rates. We maintain that all options have their individual decay rates. Because the decay rate at the syntactic level is lower than that at the morphological level, the former level increasingly facilitates the production of modals at long lags. Similarly, because the decay rate of the indicative is lower than that of the subjunctive, we observe a gradual rise in indicative options at larger distances.

Is there a connection between the activation levels in the intermediate stages and the decay rate in the final stages? It might be expected that the relative strength of a given option remains essentially the same throughout the production process. This would imply that a higher activation level correlates with a lower decay rate. Actually, the opposite is true for both the between-level and the within-level analysis. The morphological units have a higher activation level but also a higher decay rate than the modals. By the same token, the subjunctive has a higher activation level but also a higher decay rate than the indicative. This may be taken to suggest that approximately the same time span is available to all units. So when a given unit is strongly activated, it has to decay more rapidly than when it is less strongly activated.

It is possible to distinguish between a principle at work during the intermediate stages and a principle at work during the final stages. The latter is probably systemic in the sense that the linguistic system is inherently biased whereas the former is not. If there is no inherent bias, any option may predominate, as the case may be. This is what we find in the intermediate stages. While the subjunctive predominates in the modern language (see [Table 1](#)), modal constructions were the preferred choice in Middle English (Moessner 2007). Thus, the relative strength of the two options was overturned in the historical development from Middle to Modern English. A similar variation exists at the synchronic level. While some varieties of English prefer the subjunctive, others lean towards the modal construction. For example, Hundt (1998a) reports a strong predominance of the subjunctive in the Brown and Frown corpora of American English but a strong predominance of the modal construction in the Lancaster–Oslo–Bergen (LOB)

corpus and a weaker predominance of the modal construction in the Freiburg–Lancaster–Oslo–Bergen (FLOB) corpus of British English. It may be inferred that it is the individual variety which decides on the relative strength of the morphological and the syntactic options.

By contrast, it is likely that the final stages of the activation process are dominated by a processing bias which characterizes linguistic systems in general – at different times, in different varieties of the same language and possibly also in different languages. The decay rate of syntactic units is claimed to be lower than that of morphological units. Similarly, the decay rate of the indicative is claimed to be lower than that of the subjunctive. However, the reasons for the varying decay rates differ for the between-level and the within-level comparison. Let us look at each in turn, beginning with the former.

In hierarchical frameworks of language structure, syntax is superordinated to morphology. Range may straightforwardly be linked to the position of a given level in the structural hierarchy: the higher its position, the lower its decay rate and hence, the wider its range. This correlation can be turned into a cause. As the structural hierarchy is mainly motivated by the varying sizes of the elements (and their domains) at the various levels, the range of a level may be argued to be determined by the size of the units on which it operates. By virtue of their greater extension in time, larger units require wider-ranging effects for a smooth operation of the linguistic system.

In the within-level comparison, we would like to understand why the indicative has a longer range (or a lower decay rate) than the subjunctive. A possible answer can be found in the disparate use of the indicative and the subjunctive mood. Although the subjunctive has been found to be the majority option in mandative constructions (see [Table 1](#)), there is no denying that it occurs far less frequently than the indicative in general language usage. In fact, its use is highly restricted. Apart from mandatives, it occurs in optatives (e.g. *God save the Queen*), certain conditionals (e.g. Schlüter 2009), hypotheticals (e.g. Hundt, Hoffmann & Mukherjee 2012) and *lest*-clauses (e.g. Auer 2008). What is common to all these constructions is their pronounced infrequency. We therefore propose a correlation between textual frequency and range: the higher the frequency of a particular option, the longer its range.

This hypothesis appears reasonable enough. High frequency is known to increase the availability of linguistic units. It is a relatively small step from heightened availability to long range. However, this claim is in apparent conflict with syntactic persistence effects. It has been repeatedly demonstrated that less frequent constructions engender stronger persistence effects than more frequent constructions do (e.g. Bock 1986, Ferreira 2003, Rosemeyer & Schwenter 2017). We submit that persistence and mood choice represent two disparate phenomena exhibiting disparate characteristic features. Persistence involves the use of an item which was used before. That is, we are dealing with a conceptual and temporal relationship between tokens of the same type. By contrast, our study is concerned with the linear or temporal relationship between a particular trigger and

a particular target, i.e. two different words subject to morphosyntactic variation. There is another fundamental difference. In mood choice, the form of the trigger does not in any way determine the form of the target whereas persistence is defined by the influence of the first token on the likelihood of producing the second. Note further that the two phenomena have different time windows. For example, Rosemeyer & Schwenter (2017) observed persistence effects over stretches of as many as 100 words. This is much more than the (average or maximum) distance between trigger and target in our data. We thus feel justified in concluding that our frequency hypothesis is not challenged by persistence effects.

Let us briefly consider an alternative explanation of the differing ranges of the indicative and the subjunctive. Rohdenburg (1996) made a strong case for the Complexity Principle, whereby the more explicit member of a set of alternative options is preferably used in more complex contexts. This principle derives its current relevance from the fact that the indicative is explicitly coded by a suffix (in the third person singular) whereas the subjunctive is implicitly coded by the absence thereof. If the greater distance between trigger and target is taken to imply a higher level of complexity, Rohdenburg's principle correctly predicts a wider range of the indicative than of the subjunctive. One way of explaining this effect is to argue that the larger the distance between trigger and target, the higher the difficulty of establishing a link between them and hence the greater the need to facilitate the establishment of this link by means of explicit marking.

While we clearly see the merit of the Complexity Principle, we doubt that it can deal with Poplack's case of French. In contradistinction to the situation in English, it is the subjunctive rather than the indicative which is explicitly marked in many French verbs. For instance, the verbs in *-re* such as *répondre* 'to answer' have a final /d/ in the subjunctive (*qu'il réponde* /repõd/) but nothing in the indicative (*il répond* /repõd/). Still, as noted above, the relative ranges of both effects appear to be similar in the two languages. It should be added that in French, the subjunctive is not always explicitly marked and the indicative implicitly marked. Some verbs have explicit marking of both mood values while others do not formally distinguish between indicative and subjunctive mood. In either case, the Complexity Principle cannot be straightforwardly applied to French. This at least suggests the possibility that it may also not be the adequate explanation for English.

It remains for us to explain why between-level effects are stronger than within-level effects. In point of fact, this difference in strength does not come as a surprise. Given that a certain structural level has a certain range, it is to be expected that units at the same level (e.g. indicatives and subjunctives) have basically the same range while units at different levels (e.g. modals and subjunctives) differ more noticeably in range. This is just what we find. However, being located at the same level does not imply that two units must have identical ranges. There is still room for individual units to have individual ranges even though this variability is of a more limited nature.

Having dealt with range, we proffer a few comments on the effect of word class on mood choice. In both the between-level and the within-level analyses, we observed an increasing use of modal constructions and the indicative from adjectives to verbs and nouns. Actually, this is not the order which might have been expected. In many languages, adjectives take an intermediate position between verbs and nouns (e.g. Thompson 1988, Givón 1994, Stassen 1997, Pérez-Guerra 2016). Irrespective of the basis for this arrangement, such an order cannot account for our findings. What we can offer at this point is a possible correlation between word-class sensitivity and deonticity. We determined deonticity by calculating the percentages of ‘yes’ cases in the column ‘deontic use’ of our coding sheet (see Section 3) for all three word classes. The results are displayed in Table 6, which also includes the 12 imperatives in our data.

Word class	# of deontic cases	Total number	% deontic
Nouns	3,602	22,103	16.3
Verbs	9,085	19,166	48.7
Adjectives	4,427	7,385	60.0

Table 6
Deonticity of word classes.

It is apparent from Table 6 that nouns are rather reluctant to develop deontic uses. Verbs are much more likely and adjectives are most likely to do so. Thus, the three word classes can be ordered on a scale of increasing deonticity as follows: nouns > verbs > adjectives. This scale agrees to some extent with the effect of word class on mood choice. We therefore tentatively propose a partial correlation between deonticity and mood choice: the lower the degree of deonticity of a word class, the more likely it is to select a modal auxiliary or the indicative. However, this hypothesis fails to explain why nouns and verbs behave alike in our data. Such a partial correlation implies a semantic difference between the mood choices. The more specific claim would be that modals and indicatives carry a lower degree of deonticity than subjunctives and therefore combine more naturally with nouns, which rank low on the deonticity scale. This hypothesis is supported by the fact that modals code quite disparate types of meaning of which deonticity is only one.

Finally, it is fitting to return to the distinction between spoken and written language. Starting from the observation that most previous analyses of the subjunctive relied on written data, Kastronic & Poplack (2014) focus on the spoken language and report major differences between the two modes. In particular, they conclude that the subjunctive has played no more than a vestigial role in speech over the past few centuries. This conclusion is important because it challenges the received wisdom of a revival of the subjunctive in the early 20th century. Our analysis of both spoken and written samples in COCA allows us to evaluate Kastronic & Poplack’s hypothesis. It will be recalled from Section 4.2.2 above

that while there was no interaction between Distance and Mode, the subjunctive figured more prominently in writing than in speech. The actual numbers on which this finding is based are provided in [Table 7](#).

Mode	Mood		
	Indicative	Subjunctive	Total
Spoken	130 (8.2%)	1,462 (91.8%)	1,592 (100%)
Written	435 (4.9%)	8,399 (95.1%)	8,834 (100%)
Total	565	9,861	10,426

Table 7
Mood choice by mode.

In broad agreement with Kastronic & Poplack's hypothesis, the subjunctive is less frequently found in the spoken than the written data. While this difference is certainly statistically significant ($\chi^2(1) = 27.7, p < .001$), we note a difference of 'only' 3.3% between spoken and written English. This difference is not large enough to justify a radical separation of the spoken and the written language. What is more, [Table 7](#) shows that the subjunctive is vastly more frequent than the indicative in speech. Thus, there is good reason to argue that the mandative subjunctive is alive and kicking in both written and spoken American English.¹¹

6. THE WIDER PERSPECTIVE

The aim of the present section is to integrate the empirical results of this paper with previous work on distance effects. This enables us to transcend the limited scope of our study and to develop a more comprehensive model which puts our findings in a wider perspective. This leads us to a consideration of further structural levels and more radical changes in activation levels over time than were reported above.

[Figures 1 and 2](#) in [Section 4](#) showed that while the ratios increased with increasing distance, they stayed well below the value of .5. That is, the morphological level and the subjunctive mood remained the majority choice throughout. The fact that the minority option gains momentum while the majority option loses momentum over time raises the logical possibility that under appropriate circumstances we might be able to observe a cross-over effect, whereby the

[11] One reason for the disparity between Kastronic & Poplack's study and ours is that, as pointed out above, COCA's spoken sample is not truly conversational and thus more strongly influenced by the written language than Kastronic & Poplack's data are. This would suggest that the actual difference between spoken and written English may be larger than that reported in [Table 7](#).

erstwhile stronger option is eventually overridden by the erstwhile weaker option. Indeed, cross-over effects can be found in Levin (1998) for English, Leko (2000) for Bosnian and Thurmair (2006) for German. These works deal with agreement conflicts created by a competitive relationship between syntactic and semantic influences as a function of distance. Levin's (1998) analysis of agreement with collective nouns reveals that singular agreement on verbs is the rule at very short lags whereas plural agreement predominates at longer lags (see his Table 5). Similarly, Thurmair's (2006) data show that nouns with divergent natural and grammatical gender such as *das Mädchen* 'the girl' induce grammatical agreement on neighbouring targets but predominantly referential agreement on more distant targets (see Section 1 above).

These studies demonstrate that changes in effect strength over time may be more dramatic than those observed in this paper. Why do we obtain varying change rates? The major difference between the data of Levin, Leko and Thurmair and our data is the involvement of disparate structural levels. In particular, the semantic level comes into play in Levin's, Leko's and Thurmair's work but not in ours. We believe that the semantic level has a time course of activation which is rather different from that of the formal levels. To be specific, its change rate over time is relatively small. While syntactic units are gradually losing their force, semantic content persists and therefore eventually overtakes syntax. This conclusion dovetails with memory studies showing that the decay rate of semantic information is lower than that of syntactic (and other formal) information (e.g. Sachs 1967, Begg & Wickelgren 1974, Gurevich, Johnson & Goldberg 2010). Thus, the time course of activation of syntactic and morphological units may be assumed to be more similar than that of syntactic and semantic units. This is because both syntax and morphology are form-based and may perform similar jobs (see also footnote 6). On account of the relatively more similar decay rates of formal levels, it is more difficult to bring about a cross-over effect.

Prior to locating the semantic level relative to the syntactic and the morphological levels, it is fitting to bring phonology into the picture. A prime example of a syntagmatic process at the phonological level is segment or feature harmony in adult language use and child language acquisition. The range of phonological units may be determined by gauging the distance between the harmonizing and the harmonized segment. This phonological process displays a very pervasive pattern: it almost always stays within the confines of the single word. The interacting segments may, or may not, cross morpheme boundaries, but they do not usually cross word boundaries. This is true of both consonant harmony and vowel harmony (e.g. Stoel-Gammon & Stemberger 1994, van der Hulst & van der Weijer 1995, Berg 2008, Hanson 2010). It is significant in this connection that Rose & Walker (2004) define a long-distance interaction as involving segments which are only one segment apart. The few children that have been reported to harmonize consonants from different (adjacent) words (e.g. Donahue (1986)) are the exception that proves the rule. Thus, the evidence from harmony suggests that the range of phonological units roughly corresponds to a single phonological

word. Other processes such as stress assignment point to exactly the same conclusion (Kaisse 2017).

It is a straightforward task to integrate the four levels into a hierarchical framework. From the speaker's or writer's perspective, semantics is primary in that it controls the selection of linguistic units. In virtually all production models, syntax is superordinated to morphology, which in turn is superordinated to phonology. It is easy to see that the position of a level in the structural hierarchy is predictive of its range: the higher its position, the wider its range. The semantic level has the widest range and the phonological level the narrowest range. What was claimed to hold true for the syntactic and the morphological levels in the foregoing can now be claimed to hold for all four levels.

The more general view that our analysis affords of the organization of the linguistic system is one in which hierarchical position and unit size on the one hand and range and time on the other form an integrated whole. As we move from higher to lower levels, the representations gradually change. Higher levels are more strongly governed by semantic constraints while lower levels are more strongly governed by formal constraints. There are no categorical boundaries in the system. The different representational levels shade into each other. Similarly, there are no temporal boundaries. As we move from higher to lower levels, the range of linguistic units gradually diminishes. The parallelism between structure and range is all too obvious.

The agreement data also provide some insight into the initial stages of the activation process. As mentioned above, syntactic influences clearly outstrip semantic influences in agreement decisions at short lags. That is, units at the subordinate syntactic level reach a higher activation value in the initial stages than units at the superordinate semantic level. In fact, this is the prerequisite for the cross-over effect. Allowing ourselves to speculate wildly, we may generalize this claim and conjecture that the initial stages of activation witness an inverse correlation between hierarchical position and activation value. The lower the level in the hierarchy, the higher its activation value. This correlation is reversed with increasing distance. Due to differential decay rates, units at lower levels lose more of their activation than units at higher levels do in the same time frame. Thus, in the later stages of the activation process, semantic information has preserved much of its strength while phonological information has lost most or all of it. Syntactic and morphological information is in-between, with syntax losing less than morphology.

Finally, it is worthwhile comparing our account with Hawkins's (2004) Minimize Domains Principle. Essentially, the Minimize Domains Principle is about being maximally informative within a minimum amount of time. To be specific, the closer two phrase-building units are to each other, the earlier listeners can construct a syntactic representation of the input. So in case of variation, the structure with the shorter domain will be preferred to the one with the longer domain. Hawkins's theory applies (but is not logically restricted) to

the syntactic level, i.e. it deals with (within-language and between-language) syntactic variation. This is the major difference between his model and ours. Our focus is on the comparison of different levels, in particular the syntactic and the morphological. We take it that the Minimize Domains Principle applies at all relevant levels of representation. However, what our analysis shows is that the size of the domains (what we call ‘range’) is level-specific. This is why Hawkins’s model and ours are fully compatible and neatly complement each other.

7. OUTLOOK

This work aligns itself with the tradition of studies examining the impact of performance factors on the structure of grammar. It has gathered evidence showing that the choice of a grammatical marker is influenced by the linear distance between trigger and target. On the assumption that linear distance is at least a rough guide to temporal distance, it may be predicted that a grammatical decision is co-determined by the amount of time that has elapsed between the production of the trigger and that of the target. In view of the different ranges of morphological and syntactic elements, this reasoning leads to the bold prediction that a slower speech rate should encourage the use of a modal auxiliary while a faster speech rate should raise the likelihood of choosing the subjunctive.

Linear and temporal distance co-vary naturally with complexity. The greater the complexity of Y in a string XYZ, the greater the distance between X and Z. Complexity may therefore be a competitor to our distance account. Indeed, Rohdenburg (2018) has recently made a proposal along these lines. On the basis of the analysis of a single verb, he examined the use of *should* in mandative sentences as a function of the complexity of the subject of the subordinate clause and found an increased use of this modal with increasing subject complexity. From the perspective of our model, this result is to be expected. Since Rohdenburg measured complexity in terms of the number of words, complexity can be viewed as a proxy for distance. It would be interesting to pit the complexity against the distance account. Complexity is typically determined at the phrase level. Consequently, several variables (minimally, the complementizer and the subject NP) would have to be taken into account and offset against one another to bridge the gap between trigger and target verb. By contrast, the distance-based account can be carried out in one pass. Methodologically at least, it is the simpler and more elegant account.

The distance account sheds some light on an interaction which has not yet been well understood. Several studies have commented on the decrease of the modal option over the past few decades (e.g. Klein 2009, Hundt & Gardner 2017, Ruohonen 2018) while Waller (2017) observes an increasing rate of complementizer omission. Furthermore, Ruohonen (2018) reports that the rate of subjunctive use is raised by complementizer deletion. Our model not only predicts Ruohonen’s finding but also provides an explanation for it. Omitting the complementizer reduces the distance between trigger and target and hence the

likelihood of *should* productions. However, our model makes no prediction as to what is the cause and what is the effect. Exploring this connection from the diachronic perspective is a worthwhile topic for future work.

Of course, the mandative subjunctive is not the only grammatical phenomenon to span more than one level of analysis. Another area which covers both the syntactic and the morphological levels is the comparison of adjectives. English has two competing strategies of comparison – the analytic and the synthetic. The analytic technique makes use of the free-standing adverbs *more* and *most* and can therefore be located at the syntactic level. The synthetic technique relies on the suffixes *-er* and *-est* and is therefore of a morphological nature.

These disparate loci render comparison a fertile testing ground for the proposals made in this paper. Our account predicts that the range of the analytic strategy is wider than the range of the synthetic strategy. We can conceive of at least three ways in which this prediction can be put to the test.

The first analysis rests on the distinction between predication and attribution. The distance between the referent noun and the adjective differs between the predicative and the attributive use. Attributive adjectives and their nominal referents are usually adjacent whereas predicative adjectives and their referents are minimally separated by a copula. This difference allows us to predict that the analytic comparison is more frequently encountered in predicative than in attributive use while the reverse prediction holds for the synthetic comparison. This prediction was confirmed by a good number of studies (e.g. Leech & Culpeper 1997, Lindquist 2000, Mondorf 2009, Cheung & Zhang 2016).

The second analysis starts out from the disparate cohesiveness of morphemes inside words and words inside phrases. As is evidenced by the differential ease of insertion, phrases are less cohesive than words. Analytic comparatives such as *more common* may therefore be predicted to break up more easily than synthetic comparatives such as *commoner*. While this prediction may seem trivially true for synthetic comparatives, it is less obviously true for analytic comparatives. The fact that the comparative marker *more* is an independent word does not automatically imply that *more* + adjective sequences can be split. Doubly modified adjectives provide a good test case. When a graded adjective is modified by an adverb, the question arises as to where it can be put. Can it or can it not break up the *more* + adjective sequence? Behaghel's Law, according to which conceptual proximity should be mirrored by linear proximity, would lead us to expect a negative answer. By contrast, the wider range of syntactic information would provide a basis for separating the comparative marker from the adjective.

Although English is generally reluctant to tolerate discontinuity, it allows graded adjectives to be split up by so-called domain adverbs (e.g. Sullivan 2013). Pertinent examples include *more linguistically diverse* and *more cognitively complex*. Clearly, these are right-branching structures in which the domain adverb introduces discontinuity between the comparison marker and the adjective. It is a remarkable fact that this word order is the preferred choice (Berg 2019). We take the well-formedness and relative frequency of these ADJPs to be consistent

with the claim that the syntactic level has a relatively wide range. However, we are ready to acknowledge that this line of argument is of a rather more indirect nature.

Perhaps the most direct test can be achieved by a priming study. The prediction from our account is straightforward: *ceteris paribus*, analytic comparatives should prime other analytic comparatives over longer distances than synthetic comparatives should prime other synthetic comparatives. While our account makes no prediction regarding the baseline probability for suffix priming and *more*-priming, it does predict shorter distances for the former and longer distances for the latter. Such a priming experiment is a venue for future research.

The host of questions that can be raised about the comparative alternation alone convinces us that the related notions of distance and time deserve a more prominent place in the study of language than they have hitherto been assigned.

APPENDIX

Subjunctive triggers

ADJECTIVES	NOUNS	VERBS
adamant	advice	to advise
advisable	command	to advocate
anxious	condition	to arrange
appropriate	cry	to ask
better	decree	to beg
central	demand	to choose
compulsory	desire	to clamor
concerned	determination	to command
critical	dream	to decide
crucial	edict	to decree
decisive	insistence	to demand
desirable	instruction	to deserve
desirous	mandate	to desire
determined	motion	to determine
eager	plan	to dictate
elementary	principle	to direct
essential	priority	to ensure
expedient	proposal	to expect
fair	proposition	to insist
fitting	provision	to intimate
fundamental	recommendation	to make sure
imperative	remedy	to move
important	request	to ordain
indispensable	requirement	to order
insistent	resolution	to persuade
integral	restriction	to petition
intent	rule	to plead

ADJECTIVES	NOUNS	VERBS
keen	ruling	to prefer
key	stipulation	to propose
mandatory	suggestion	to provide
natural	supplication	to recommend
necessary	understanding	to request
obligatory	urge	to require
paramount	will	to resolve
pivotal		to see to it
preferable		to specify
proper		to stipulate
prudent		to suggest
requisite		to take care
undesirable		to urge
urgent		to want
vital		to will
willing		to wish

REFERENCES

- Arnold, Jennifer E., Thomas Wasow, Anthony Losongco & Ryan Ginstrom. 2000. Heaviness vs. newness: The effects of structural complexity and discourse status on constituent ordering. *Language* 76, 28–55.
- Auer, Anita. 2008. *Lest* the situation deteriorates: A study of *lest* as trigger of the inflectional subjunctive. In Miriam A. Locher & Jürg Strässler (eds.), *Standards and norms in the English language*, 149–173. Berlin: Mouton de Gruyter.
- Barr, Dale J., Roger Levy, Christoph Scheepers & Harry J. Tily. 2013. Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language* 68, 255–278.
- Bates, Douglas, Martin Maechler, Ben Bolker & Steve Walker. 2015. Fitting linear mixed-effects models using *lme4*. *Journal of Statistical Software* 67, 1–48.
- Begg, Ian & Wayne A. Wickelgren. 1974. Retention functions for syntactic and lexical vs. semantic information in sentence recognition memory. *Memory & Cognition* 2, 353–359.
- Behaghel, Otto. 1909. Beziehungen zwischen Umfang und Reihenfolge von Satzgliedern. *Indogermanische Forschungen* 25, 110–142.
- Belnap, R. Kirk. 1999. A new perspective on the history of Arabic variation in marking agreement with plural heads. *Folia Linguistica* 33, 169–185.
- Berg, Thomas. 2008. On the complementary nature of consonant and vowel harmony in child and adult language. *Word* 59, 241–256.
- Berg, Thomas. 2019. Adjective phrases with doubly modified heads: How lexical information influences word order and constituent structure. *English Language & Linguistics* 23, 341–361.
- Berlage, Eva. 2014. *Noun phrase complexity in English*. Cambridge: Cambridge University Press.
- Bock, J. Kathryn. 1986. Syntactic persistence in language production. *Cognitive Psychology* 18, 355–387.
- Bock, J. Kathryn & J. Cooper Cutting. 1992. Regulating mental energy: Performance units in language production. *Journal of Memory and Language* 31, 99–127.
- Branigan, Holly P., Martin J. Pickering & Alexandra A. Cleland. 1999. Syntactic priming in written production: Evidence for rapid decay. *Psychonomic Bulletin & Review* 6, 633–640.
- Bresnan, Joan, Anna Cueni, Tatiana Nikitina & R. Harald Baayen. 2007. Predicting the dative alternation. In Gerlof Bouma, Irene Krämer & Joost Zwarts (eds.), *Cognitive foundations of interpretation*, 69–94. Amsterdam: Royal Netherlands Academy of Arts and Sciences.
- Bybee, Joan L. 1985. *Morphology*. Amsterdam: John Benjamins.

- Bybee, Joan L., Revere Perkins & William Pagliuca. 1994. *The evolution of grammar*. Chicago, IL: University of Chicago Press.
- Cheung, Lawrence & Longtu Zhang. 2016. Determinants of the synthetic: Analytic variation across English comparatives and superlatives. *English Language & Linguistics* 20, 559–583.
- Clogg, Clifford C., Eva Petkova & Adamantios Haritou. 1995. Statistical methods for comparing regression coefficients between models. *American Journal of Sociology* 100, 1261–1293.
- Corbett, Greville G. 1991. *Gender*. Cambridge: Cambridge University Press.
- Crawford, William J. 2009. The mandative subjunctive. In Günter Rohdenburg & Julia Schlüter (eds.), *One language, two grammars? Differences between British and American English*, 257–276. Cambridge: Cambridge University Press.
- Curzan, Anne. 2003. *Gender shifts in the history of English*. Cambridge: Cambridge University Press.
- Donahue, Mavis. 1986. Phonological constraints on the emergence of two-word utterances. *Journal of Child Language* 13, 209–218.
- Erdmann, Peter. 1981. Der Konjunktiv im britischen und amerikanischen Englisch. In Peter Kunsman & Ortwin Kuhn (eds.), *Weltsprache Englisch in Forschung und Lehre*, 110–131. Berlin: Schmidt.
- Ferreira, Fernanda. 1991. Effects of length and syntactic complexity on initiation times for prepared utterances. *Journal of Memory and Language* 30, 210–233.
- Ferreira, Victor S. 2003. The persistence of optional complementizer production: Why saying *that* is not saying *that* at all. *Journal of Memory and Language* 48, 379–398.
- Futrell, Richard, Kyle Mahowald & Edwin Gibson. 2015. Large-scale evidence of dependency length minimization in 37 languages. *Proceedings of the National Academy of Sciences* 112, 10336–10341.
- Gibson, Edward. 2000. The Dependency Locality Theory: A distance-based theory of linguistic complexity. In Alec Marantz, Yasushi Miyashita & Wayne O’Neil (eds.), *Image, language, brain*, 95–126. Cambridge, MA: MIT Press.
- Gildea, Daniel & David Temperley. 2010. Do grammars minimize dependency length? *Cognitive Science* 34, 286–310.
- Gillespie, Maureen & Neal J. Perlmuter. 2011. Hierarchy and scope of planning in subject–verb agreement production. *Cognition* 118, 377–397.
- Givón, Talmay. 1994. *Syntax: A functional-typological introduction*. Amsterdam: John Benjamins.
- Greenbaum, Sidney & Randolph Quirk. 1990. *A student’s grammar of the English language*. Harlow: Longman.
- Gries, Stefan Th. 2002. Preposition stranding in English: Predicting speakers’ behaviour. In Vida Samiian (ed.), *Proceedings of the Western Conference on Linguistics*, vol. XII (WECOL 12), 230–241. Fresno, CA: Department of Linguistics, California State University.
- Gries, Stefan Th. 2003. *Multifactorial analysis in corpus linguistics: A study of particle placement*. London: Continuum.
- Gries, Stefan Th. 2005. Syntactic priming: A corpus-based approach. *Journal of Psycholinguistic Research* 34, 365–399.
- Gries, Stefan Th. & Martin Hilpert. 2010. Modeling diachronic change in the third person singular: A multifactorial, verb- and author-specific exploratory approach. *English Language & Linguistics* 14, 293–320.
- Gurevich, Olga, Matthew A. Johnson & Adele E. Goldberg. 2010. Incidental verbatim memory for language. *Language and Cognition* 2, 45–78.
- Hanson, Gunnar Ólafur. 2010. *Consonant harmony: Long-distance interaction in phonology*. Berkeley, CA: University of California Press.
- Haspelmath, Martin. 2011. The indeterminacy of word segmentation and the nature of morphology and syntax. *Folia Linguistica* 45, 31–80.
- Hawkins, John A. 1994. *A performance theory of order and constituency*. Cambridge: Cambridge University Press.
- Hawkins, John A. 1999. The relative order of prepositional phrases in English: Going beyond manner–place–time. *Language Variation and Change* 11, 231–266.
- Hawkins, John A. 2004. *Efficiency and complexity in grammars*. Oxford: Oxford University Press.
- Hay, Jennifer. 2001. Lexical frequency in morphology: Is everything relative? *Linguistics* 39, 1041–1070.

- Hilpert, Martin. 2008. The English comparative: Language structure and language use. *English Language & Linguistics* 12, 395–417.
- Hinrichs, Lars, Benedikt Szmrecsanyi & Axel Bohmann. 2015. Which-hunting and the Standard English relative clause. *Language* 91, 806–836.
- Hoffmann, Sebastian. 1997. Mandative sentences: A study of variation on the basis of the British National Corpus. MA thesis, University of Zurich.
- Hoffmann, Thomas. 2011. *Preposition placement in English: A usage-based approach*. Cambridge: Cambridge University Press.
- Horn, George M. 1986. The minimal distance principle revisited. In Dieter Kastovsky & Aleksander Szwedek (eds.), *Linguistics across historical and geographical boundaries*, vol. 2, 909–934. Berlin: Mouton de Gruyter.
- Huddleston, Rodney. 2002. Content clauses and reported speech. In Rodney Huddleston & Geoffrey K. Pullum et al. *The Cambridge grammar of the English language*, 947–1030. Cambridge: Cambridge University Press.
- Hundt, Marianne. 1998a. It is important that this study (*should*) be based on the analysis of parallel corpora: On the use of the mandative subjunctive in four major varieties of English. In Hans Lindquist, Staffan Klintberg, Magnus Levin & Maria Estling (eds.), *The major varieties of English*, 159–175. Växjö: Acta Wexioniensia.
- Hundt, Marianne. 1998b. *New Zealand English grammar: Fact or fiction*. Amsterdam: John Benjamins.
- Hundt, Marianne & Anne-Christine Gardner. 2017. Corpus-based approaches: Watching English change. In Laurel Brinton (ed.), *English historical linguistics*, 96–130. Cambridge: Cambridge University Press.
- Hundt, Marianne, Sebastian Hoffmann & Joybrato Mukherjee. 2012. The hypothetical subjunctive in South Asian Englishes. *English World-Wide* 33, 147–164.
- Jaeger, T. Florian & Neal Snider. 2008. Implicit learning and syntactic persistence: Surprisal and cumulativity. *Proceedings of the 30th Annual Conference of the Cognitive Science Society*, 1061–1066.
- Jeschull, Liane. 2004. Coordination in Chechen. In Martin Haspelmath (ed.), *Coordinating constructions*, 241–265. Amsterdam: John Benjamins.
- Johansson, Stig & Else Helene Norheim. 1988. The subjunctive in British and American English. *ICAME Journal* 12, 27–36.
- Kaisse, Ellen. 2017. The domain of stress assignment: Word-boundedness and frequent collocation. In Claire Bowers, Lawrence Horn & Raffaella Zanuttini (eds.), *On looking into words (and beyond)*, 17–40. Berlin: Language Science Press.
- Kastronic, Laura & 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, Article 9. <http://www.sociolinguistics.uottawa.ca/shanapoplack/pubs/articles/KastronicPoplack2014.pdf> (accessed 21 March 2019).
- Klein, Barbara. 2009. Ongoing grammatical change in spoken British English: Real-time studies based on the DCPSE. *Arbeiten aus Anglistik und Amerikanistik* 34, 29–48.
- Köpcke, Klaus-Michael, Klaus-Uwe Panther & David Zubin. 2009. Motivating grammatical and conceptual agreement in German. In Hans-Jörg Schmid & Susanne Handl (eds.), *Cognitive foundations of linguistic usage patterns*, 171–194. Berlin: Mouton de Gruyter.
- Leech, Geoffrey & Jonathan Culpeper. 1997. The comparison of adjectives in recent British English. In Terttu Nevalainen & Leena Kahlas-Tarkka (eds.), *To explain the present: Studies in the changing English language in honour of Matti Rissanen*, 353–373. Helsinki: Société Néophilologique.
- Leech, Geoffrey, Marianne Hundt, Christian Mair & Nicholas Smith. 2009. *Change in contemporary English: A grammatical study*. Cambridge: Cambridge University Press.
- Leko, Nedžad. 2000. Syntactic vs. semantic agreement in the Oslo Corpus of Bosnian texts. In Tracy Holloway King & Irina A. Sekerina (eds.), *Annual Workshop on Formal Approaches to Slavic Linguistics 8: The Philadelphia Meeting 1999*, 259–278. Ann Arbor, MI: Michigan Slavic Publications.
- Levin, Magnus. 1998. On concord with collective nouns in English. In Antoinette Renouf (ed.), *Explorations in corpus linguistics*, 99–112. Amsterdam: Rodopi.
- Levin, Magnus. 2001. *Agreement with collective nouns in English*. Stockholm: Almqvist & Wiksell.

- Lindquist, Hans. 2000. *Livelier or more lively?* Syntactic and contextual factors influencing the comparison of disyllabic adjectives. In John M. Kirk (ed.), *Corpora galore: Analyses and techniques in describing English*, 125–132. Amsterdam: John Benjamins.
- Lohmann, Arne & Tayo Takada. 2014. Order in NP conjuncts in spoken English and Japanese. *Lingua* 152, 48–64.
- Lohse, Barbara, John A. Hawkins & Thomas Wasow. 2004. Domain minimization in English verb-particle constructions. *Language* 80, 238–261.
- Marslen-Wilson, William, Lorraine K. Tyler, Rochelle Waksler & Lianne Older. 1994. Morphology and meaning in the English mental lexicon. *Psychological Review* 101, 3–33.
- Moessner, Lilo. 2007. The mandative subjunctive in Middle English. In Gabriela Mazzon (ed.), *Studies in Middle English forms and meanings*, 209–226. Frankfurt a.M.: Lang.
- Mondorf, Britta. 2009. *More support for more-support*. Amsterdam: John Benjamins.
- Morgan, James L. 1972. Verb agreement as a rule of English. In Paul M. Peranteau, Judith N. Levi & Gloria C. Phares (eds.), *Papers from the 8th Regional Meeting of the Chicago Linguistic Society* (CLS 8), 278–286. Chicago, IL: Chicago Linguistic Society.
- Övergaard, Gerd. 1995. *The mandative subjunctive in American and British English in the 20th century*. Uppsala: Almqvist & Wiksell.
- Pérez-Guerra, Javier. 2016. Do you investigate word order in detail or do you investigate in detail word order? On word order and headedness in the recent history of English. *Corpus Linguistics and Linguistic Theory* 12, 103–128.
- Peters, Pam. 2009. The mandative subjunctive in spoken English. In Pam Peters, Peter Collins & Adam Smith (eds.), *Comparative studies in Australian and New Zealand English*, 125–137. Amsterdam: John Benjamins.
- Pietsch, Lukas. 2005. *Variable grammars: Verbal agreement in northern dialects of English*. Tübingen: Niemeyer.
- Plank, Frans. 1985. Die Ordnung der Personen. *Folia Linguistica* 19, 111–176.
- Poplack, Shana. 1992. The inherent variability of the French subjunctive. In Christiane Laeuffer & Terrell A. Morgan (eds.), *Theoretical analyses in Romance linguistics*, 235–263. Amsterdam: John Benjamins.
- Poplack, Shana, Allison Lealess & Nathalie Dion. 2013. The evolving grammar of the French subjunctive. *Probus* 25, 139–195.
- R Development Core Team. 2014. R: A language and environment for statistical computing. <http://cran.r-project.org/>.
- Reitter, David, Johanna D. Moore & Frank Keller. 2006. Priming of syntactic rules in task-oriented dialogue and spontaneous conversation. *Proceedings of the 28th Annual Conference of the Cognitive Science Society*, 1–6.
- Roelofs, Ardi. 2002. Syllable structure effects turn out to be word length effects: Comment on Santiago et al. (2000). *Language and Cognitive Processes* 17, 1–13.
- Rohdenburg, Günter. 1996. Cognitive complexity and increased grammatical explicitness in English. *Cognitive Linguistics* 7, 149–182.
- Rohdenburg, Günter. 1999. Clausal complementation and cognitive complexity in English. In Fritz-Wilhelm Neumann & Sabine Schülting (eds.), *Anglistentag 1998*, 101–112. Trier: Wissenschaftlicher Verlag.
- Rohdenburg, Günter. 2018. The use of optional complement markers in present-day English. In Mark Kaunisto, Mikko Höglund & Paul Rickman (eds.), *Changing structures: Studies in construction and complementation*, 129–149. Amsterdam: John Benjamins.
- Rose, Sharon & Rachel Walker. 2004. A typology of consonant agreement as correspondence. *Language* 80, 475–531.
- Rosemeyer, Malte & Scott A. Schwenter. 2017. Entrenchment and persistence in language change: The Spanish past subjunctive. *Corpus Linguistics and Linguistic Theory* 13, 1–38.
- Rosenbaum, Peter S. 1970. A principle governing deletion in English sentential complementation. In Roderick A. Jacobs & Peter S. Rosenbaum (eds.), *Readings in English transformational grammar*, 20–29. Waltham, MA: Ginn.
- Ruohonen, Juko. 2018. 21st-century trends in BrE mandative constructions: A statistical multivariate analysis. Presented at the 5th Conference of the International Society for the Linguistics of English, London.

- Sachs, Jacqueline S. 1967. Recognition memory for syntactic and semantic aspects of connected discourse. *Perception & Psychophysics* 2, 437–442.
- Santiago, Julio, Donald G. MacKay & Alfonso Palma. 2002. Length effects turn out to be syllable structure effects: Response to Roelofs (2002). *Language and Cognitive Processes* 17, 15–29.
- Santiago, Julio, Donald G. MacKay, Alfonso Palma & Christine Rho. 2000. Sequential activation processes in producing words and syllables: Evidence from picture naming. *Language and Cognitive Processes* 15, 1–44.
- Schlüter, Julia. 2009. The conditional subjunctive. In Günter Rohdenburg & Julia Schlüter (eds.), *One language, two grammars? Differences between British and American English*, 277–305. Cambridge: Cambridge University Press.
- Schneider, Edgar W. 2005. The subjunctive in Philippine English. In Danilo T. Dayag & J. Stephen Quakenbush (eds.), *Linguistics and language education in the Philippines and beyond*, 27–40. Manila: Linguistic Society of the Philippines.
- Serpellet, Noelle. 2001. The mandative subjunctive in British English seems to be alive and kicking. . . Is this due to the influence of American English? In Paul Rayson, Andrew Wilson, Tony McEnery, Andrew Hardie & Shereen Khoja (eds.), *Proceedings of the Corpus Linguistics 2001 Conference*, 531–542. Lancaster: University Centre for Computer Corpus Research on Language.
- Sridhar, Shikaripur N. 1990. *Kannada*. London: Routledge.
- Stassen, Leon. 1997. *Intransitive predication*. Oxford: Clarendon.
- Stoel-Gammon, Carol & Joseph P. Stemberger. 1994. Consonant harmony and phonological underspecification in child speech. In Mehmet Yavas (ed.), *First and second language phonology*, 63–80. San Diego, CA: Singular Publishing.
- Sullivan, Keren. 2013. *Frames and constructions in metaphorical language*. Amsterdam: John Benjamins.
- Swinney, David A. 1979. Lexical access during sentence comprehension: (Re)consideration of context effects. *Journal of Verbal Learning and Verbal Behavior* 18, 645–659.
- Szmrecsanyi, Benedikt M. 2004. On operationalizing syntactic complexity. In Gérard Purnelle, Cédric Fairon & Anne Dister (eds.), *Le poids des mots*, 1032–1039. Louvain-La-Neuve: Presses Universitaires de Louvain.
- Szmrecsanyi, Benedikt M. 2006. *Morphosyntactic persistence in spoken English: A corpus study at the intersection of variationist sociolinguistics, psycholinguistics, and discourse analysis*. Berlin: Mouton de Gruyter.
- Szmrecsanyi, Benedikt M., Jason Grafmiller, Benedikt Heller & Melanie Röthlisberger. 2016. Around the world in three alternations: Modeling syntactic variation in varieties of English. *English World-Wide* 27, 109–137.
- Tarallo, Fernando. 1986. Functional and structural properties in a variable system. In David Sankoff (ed.), *Diversity and diachrony*, 249–260. Amsterdam: John Benjamins.
- Thompson, Sandra A. 1988. A discourse approach to the cross-linguistic category ‘adjective’. In John A. Hawkins (ed.), *Explaining language universals*, 167–185. London: Blackwell.
- Thurmair, Maria. 2006. *Das Modell und ihr Prinz*. Kongruenz und Texteinbettung bei Genus-Sexus-Divergenz. *Deutsche Sprache* 34, 191–220.
- Trotta, Joe. 2000. *Wh-clauses in English: Aspects of theory and description*. Amsterdam: Rodopi.
- van der Hulst, Harry & Jeroen van der Weijer. 1995. Vowel harmony. In John Goldsmith (ed.), *The handbook of phonological theory*, 495–534. Oxford: Blackwell.
- Vigliocco, Gabriella & Janet Nicol. 1998. Separating hierarchical relations and word order in language production: Is proximity concord syntactic or linear? *Cognition* 68, B13–B29.
- Vlasova, Olga. 2010. The mandative subjunctive in American English. MA thesis, University of Oslo.
- Waller, Tim. 2017. *The subjunctive in present-day English*. Ph.D. thesis, University College London.
- Wasow, Thomas. 1997. Remarks on grammatical weight. *Language Variation and Change* 9, 81–105.
- Wasow, Thomas & Jennifer Arnold. 2003. Post-verbal constituent ordering in English. In Günter Rohdenburg & Britta Mondorf (eds.), *Determinants of grammatical variation in English*, 119–154. Berlin: Mouton de Gruyter.
- Wiechmann, Daniel & Arne Lohmann. 2013. Domain minimization and beyond: Modelling prepositional phrase ordering. *Language Variation and Change* 25, 65–88.
- Yaruss, J. Scott. 1999. Utterance length, syntactic complexity, and childhood stuttering. *Journal of Speech, Language, and Hearing Research* 42, 329–344.

Authors' addresses: (Berg)

*Department of English, University of Hamburg,
Überseering 35, 22297 Hamburg, Germany
thomas_berg@uni-hamburg.de*

(Zingler)

*Department of Linguistics, University of New Mexico,
Albuquerque, New Mexico 87131, USA
timzingler@unm.edu*

(Lohmann)

*Department of English, Heinrich-Heine-Universität Düsseldorf,
Universitätsstrasse 1, 40225 Düsseldorf, Germany
Arne.Lohmann@uni-duesseldorf.de*