

Although-constructions in varieties of English

Ole Schützler

English Linguistics, University of Bamberg,
Germany

Correspondence

Ole Schützler, English Linguistics, University of
Bamberg, Bamberg, Germany.

Email: ole.schuetzler@uni-bamberg.de

Abstract

Focusing on subordinate clauses of concession introduced by *although*, this paper investigates the systematic variation of clause position by inspecting three language-internal factors (or constraints) and their stability across different varieties of the language. The three constraints are based on production (with speech making final placement more likely, due to reduced planning load), processing (with a general preference of final position), and semantics (with certain intra-constructural relations favouring initial placement due to the iconicity principle). The strength and robustness of the constraints outlined above are tested based on spoken and written corpus data from British, Canadian, New Zealand, Nigerian, Indian, and Philippine English.

1 | INTRODUCTION

This paper explores syntactic variation in concessive constructions with the subordinator *although*. The focus is on three constraints: (i) a production-based constraint whereby the higher time pressure in speech (vs. writing) makes advance planning more difficult and results in the final placement of subordinate clauses; (ii) a processing-based constraint whereby clauses are easier to parse and therefore preferred in final position; and (iii) a semantic constraint, whereby certain intra-constructural relationships between propositions will promote arrangements of clauses that are iconic of those semantic relationships. Including data from several varieties of English opens up another theoretical dimension: by showing whether or not the constraints outlined above are uniform across L1 and L2 varieties of English, we can explore whether they are based upon truly universal, cognitive processes (in which case there would be no difference between L1 and L2 Englishes), or whether they are affected by long-term nativization processes. The study approaches these issues in a quantitative manner, drawing on data from different national components of the *International Corpus of English* (ICE; Greenbaum, 1992; Nelson, 1996). Comparisons are drawn between English in Great Britain (BrE), Canada (CanE), New Zealand (NZE), Nigeria (NigE), India (IndE), and the Philippines (PhiE).

While there have been many theoretical and some quantitative studies of the semantics of concessives (Azar, 1997; Di Meola, 1998; Hermodsson, 1994; Hilpert, 2013; König, 1988, 1991, 2006), the variability of clause position has been

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2020 The Authors. *World Englishes* published by John Wiley & Sons Ltd

less studied (Altenberg, 1986; Wiechmann & Kerz, 2013). This paper adds to the existing research by approaching concessive constructions from the perspective of world Englishes, and by explicitly inspecting the different constraints outlined above. Section 2 outlines the two dimensions of variability relevant for the concessive constructions discussed in this paper: (i) the semantic (or pragmatic) relationship between the two propositions that make up a construction and (ii) the arrangement of subordinate clauses relative to matrix clauses. Furthermore, it presents some basic descriptive statistics. Section 3 summarizes the more general aspects that are used in hypothesis-building: speech vs. writing, production, processing, iconicity, and the potential effects of variety-status. Section 4 briefly discusses issues of data retrieval, annotation, and statistical analysis. Section 5 provides the main analysis and results. Finally, Section 6 presents a brief conclusion and outlook to future research.

2 | THE VARIABILITY OF CONCESSIVE CONSTRUCTIONS

The variability of concessive constructions is discussed in terms of two dimensions: (i) the intra-constructional semantic/pragmatic make-up of the concessive; and (ii) the syntactic sequencing of matrix and subordinate clause. As an additional dimension of variability, the clause-internal syntax of the subordinate clause (for example finite vs. non-finite/reduced) could also be inspected, but this was not included in the present study. Finally, there is the issue of the (relative) weight or complexity of matrix and subordinate clauses, whose correlation with clause position has been demonstrated numerous times and is relatively uncontroversial, with longer/heavier structures tending to take final position (Diessel, 2005; Schützler, 2018; Wiechmann & Kerz, 2013). However, it is perhaps less straightforward to construe length as a conditioning factor in this context, as it seems doubtful that the language user has an idea of precisely how much he or she wants to say before the utterance is initiated – a situation that is of course different in writing. A more convincing interpretation would be that non-final position acts as a brake on excessively long subordinate clauses.¹ The focus in this study is, therefore, on the mode of production and the intra-constructional semantics, since causes and effects are less easily confused with regard to these predictors: the context of production (spoken vs. written) is a given, and with regard to semantics, I assume that form follows function.

2.1 | Semantic and pragmatic relations

Using a slightly different terminology, this paper adopts Sweetser's (1990) framework of three types of concessives (Crevels, 2000; Hilpert, 2013). 'Anticausal concessives' (called 'content concessives' by Sweetser) are based on some causal or conditional presupposition, whose non-realisation results in concessive meaning. Thus, in contrast to the facts presented in (1), a newly assembled team would normally be expected to take a relatively long time to develop a harmonic workflow.

- (1) **Although** this team has not been together very long, they are developing a nice rapport between themselves [...]. (ICE-CAN:S2A-014)

The concessive relationship between propositions is based on some implied (and generally understood) mechanism of real-world causality or conditionality, which Anscombe (1989) refers to as a 'topos'. In this case the topos could be summarized as 'team-building takes time'.

Concessives of the second type, 'epistemic concessives' (which Sweetser handles under the same label), trigger an inference that turns out to be false. Thus, in (2) a person regularly posing with a certain weapon – in portraits, we must assume – is expected to do so because he actually fights (or fought) with it, too.

- (2) [...A]**lthough** he's always brandishing his bolo, Bonifacio never fought with the bolo. (ICE-PHI:S2A-034)

The difference between anticausal and epistemic concessives is that the former involves real-world conditionality, while the latter does not. Based on the propositional content of the subordinate clause in an anticausal concessive, we expect some outcome, which then remains unrealized. In an epistemic concessive, we observe an outcome and draw certain conclusions as to the underlying causes: in (2), Bonifacio's brandishing a bolo in portraits of himself cannot be construed as the cause of having fought with this kind of weapon. In a sense, the direction of inference is inverted from cause → effect in anticausal concessives to effect → cause in epistemic concessives.

Finally, in 'dialogic concessives' one proposition pragmatically qualifies or corrects another, as in (3), where positive future prospects (and their implications) are juxtaposed with a less optimistic present situation, with no clear causal or inferential trajectory between the two stances.

(3) I hope to raise the money for the purpose, **although** at present I have not got anything. (ICE-IND:W1B-012)

In the context of the present study, the notions of 'inference', 'conditionality', or 'cause and effect' are helpful, since they imply a semantic directionality that will play a role in the formulation of hypotheses in Section 3.4 below.²

2.2 | Clause positions

Subordinate adverbial clauses may occur in initial, medial, and final position (Huddleston & Pullum, 2002, p. 779; Quirk, Greenbaum, Leech, & Svartvik, 1985, p. 1037). Elements in final position are placed after the matrix-clause verb and its complements/objects; elements in medial position are placed after the subject but before the verb; and elements in initial position precede the matrix-clause subject (Huddleston & Pullum, 2002, p. 780). All three realisations are shown in (4).

(4) a. Patience was [...] already greyer-haired than Miriam, **although** she was eleven years her junior. (ICE-GB:W2F-007)

b. Patience, **although** she was eleven years her sister's junior, was already greyer-haired than Miriam. (reconstructed from 4a)

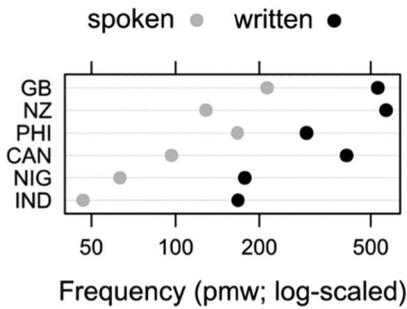
c. **Although** she was eleven years her sister's junior, Patience was already greyer-haired than Miriam. (reconstructed from 4a)

There is no obvious difference in meaning between the three variant constructions, but the information is structured differently, since clausal material in final position tends to be given greater emphasis, according to Quirk et al.'s (1985, p. 1036) principle of resolution. However, as discussed in the following section, this study draws on other explanations for positional variation in complex sentences.

2.3 | Descriptive statistics

Figure 1 summarizes some basic characteristics of clauses with *although*. On the left, normalized frequencies of *although* in speech and writing are shown for the six varieties under investigation. On the right, proportions of the three semantic types are shown. In this case, there is no differentiation of varieties, since values are very similar in all of them, which is also indicated by the rather narrow 50 per cent and 89 per cent uncertainty intervals.³ Clauses with *although* are considerably more frequent in writing than in speech, a pattern also typical of other (concessive) subordinators (Schützler, 2017, p. 178). As shown in Figure 1a, all six varieties behave similarly in this regard, even if in some varieties (at the bottom of the panel) the overall frequency of *although* is relatively low. Concerning typical semantic structures, Figure 1b shows that epistemic concessives are rare, while dialogic concessives are the majority type. There is but a small difference between speech and writing in this regard. Due to its loosely qualifying, quasi-coordinated character (see 2.1 above), the dialogic type might have been expected to occur at higher rates in speech; this, however, is not borne out by the data. Note that estimates for all three proportions are very robust, even though only $n = 6$ varieties

(a) Frequencies



(b) Semantic types

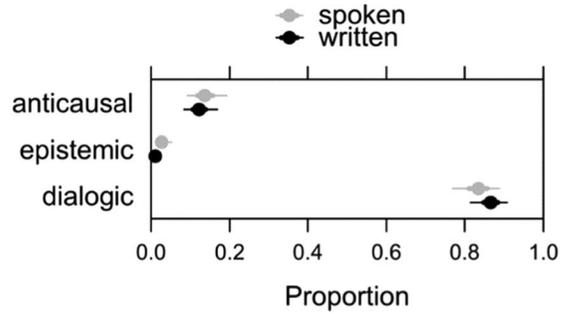


FIGURE 1 Frequencies of *although* (a) and proportions of semantic types (b)

were sampled at the highest level of the hierarchical data structure – that is, the pattern is highly homogeneous across varieties (see full cross-tabulation of data in Appendix A). Due to their rarity, epistemic concessives are excluded from the analyses in Section 5.

3 | THEORETICAL CONSIDERATIONS

In Section 3.1, production and processing are discussed together with general aspects concerning the difference between speech and writing, since these issues are regarded as interlinked; Section 3.2 applies the iconicity principle to adverbial clauses of concession; Section 3.3 outlines some assumptions about (different classes of) world Englishes and how they can inform expectations in this particular study; and Section 3.4 formulates four hypotheses that will be put to the quantitative test in Section 5.

3.1 | Speech and writing, production and processing

Prototypical speech is evanescent, relatively quickly and spontaneously produced, and clearly situated concerning place, time, and interlocutors. Writing, on the other hand, is permanent, produced more slowly than speech, and may be edited and revised; furthermore, it is desituated, that is, less clearly tied to a particular temporal, local, or circumstantial context (Chafe, 1994, pp. 42–45; Linell, 2005, p. 21). In consequence, speech and writing differ with regard to planning, production, and processing: a speaker has less time to think ahead and anticipate complex structures than a writer and thus will tend to produce structures that are smaller and more hermetic, rather than embedded in larger complex structures. Similarly, a listener is subject to more rigid constraints than a reader; the latter can re-read and make sense of extremely complex passages, while the former is constantly confronted with new material to process and thus has little opportunity to re-interpret earlier structures. Finally, there is also a crucial difference between speech and writing in language acquisition, which can help account for differences in the use of certain constructions in the two modes of production. As Akinaso (1982, p. 111) points out, speech is for the most part acquired ‘naturally’, not at school. The same point is made by Linell (2005, p. 23), who argues that more explicit instruction and norms are involved in learning to write. The presence (or absence) of certain norms concerning the use of concessives in works of reference may contribute to explanations of patterns found particularly in L2 varieties, in which English is acquired scholastically to a greater extent.

Chafe (1985, p. 111; Biber, 1988; Chafe & Danielewicz, 1987, p. 21) also discusses the difference between speech and writing in terms of fragmentation vs. integration. In more fragmented (spoken) discourse, ‘idea units tend to be

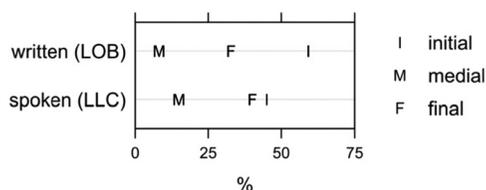


FIGURE 2 Positions of *although*-clauses in writing and speech (Altenberg, 1986)

strung together in a chain, with a relatively small amount of subordination' (Chafe, 1985, p. 111; Pawley & Syder, 1983, pp. 565–567). This notion can be extended to explain the positional variation of concessive clauses: if a subordinate clause precedes the matrix clause, it will tend to hold the entire construction in suspense, since the semantic processing of the whole is possible only after production of the matrix clause. If, on the other hand, the subordinate clause follows the matrix clause, the latter will be more self-sufficient and can be processed more locally, even if its interpretation is modified by the following subordinate clause (Huddleston & Pullum, 2002, p. 780; Quirk et al., 1985, pp. 1039–1040). The sequence of 'matrix → subordinate' is therefore more chain-like (in the sense of Chafe, 1985) than the inverse arrangement, and would be expected to be relatively more frequent in speech. This argumentation ties in with the theory proposed by Hawkins (1990, 1994; see below), which is more exclusively motivated from considerations of syntactic processing (parsing), however. Very similarly, Altenberg (1986, p. 21) argues that a preposed subordinate clause has a 'grounding' function – that is, it provides background information against which the (more important) information in the main clause can be evaluated. According to Altenberg, this arrangement of 'subordinate → matrix' entails more rigorous advance planning on the part of the speaker or writer, which is why sentence-initial subordinate clauses are more likely in the written mode. By contrast, conversation is characterised by planning in 'real time', with 'locally managed' units (again, note the correspondence to notions like 'chain-like'): '[W]hen planning is not far ahead of production, it is easier to qualify a superordinate idea retrospectively (by postposition) than to anticipate it by means of grounding (pre-position)' (Altenberg, 1986, p. 21). Altenberg present a quantitative analysis comparing clause positions in the *London-Lund Corpus of Spoken English* (LLC; Svartvik & Quirk, 1980) and the (written) *Lancaster-Oslo/Bergen Corpus* (LOB; Johansson, Leech, & Goodluck, 1978). While clauses with *although* generally tend to appear in initial position in Altenberg's data, the difference between speech and writing is nevertheless clearly reflected, as shown in Figure 2. The relative frequency of initial subordinate *although*-clauses is higher in writing, while the relative frequency of subordinate clauses in medial and final position is higher in speech.⁴

Relying on Hawkins's (1990, 1994) 'performance theory of order and constituency', Diessel (2005, pp. 458–459) argues that the preferred position for adverbial clauses is sentence-final, from the perspectives of both production and parsing.⁵ Since the matrix clause is constructed first in this case, the language producer does not need to make an early commitment to a complex sentence structure and is thus relieved of advance planning. And, since the subordinating conjunction follows the matrix clause, it marks the entire sentence as complex and indicates the boundary between matrix and subordinate clause at the same time. Finally, no (or at least considerably weaker) weight-constraints are placed upon a subordinate clause in final position. On this basis, the non-final placement of subordinate clauses appears as the marked solution which needs to be motivated, for example invoking iconicity as discussed in the following section. To return to the distinction between speech and writing above, it can easily be imagined that predictions as made by Hawkins (1990, 1994) will apply less to written language.

3.2 | Iconicity

Crystal (1992, p. 179) defines iconicity as '[a] close physical relationship between a linguistic sign [...] and the entity or process in the world to which it refers'. The iconicity principle runs counter to the basic structuralist view that linguistic signs are arbitrary (Simone, 1995, p. 157). We acknowledge that linguistic structures (and thus the linguistic

system) are not entirely independent but potentially motivated by the conceptual structures for which they stand; these conceptual structures can then to some extent account for (or functionally explain) the chosen form (Haiman, 1985, p. 1). Word order iconicity is one of the most recognized and best documented types, the idea being that 'the order of clauses or phrases corresponds to the order of our thoughts' (Haiman, 2003), which are mostly based on real-world mechanisms and expectations.⁶ Simone (1995, p. 161) cites the example reproduced as (5) and argues that the two imperative clauses can only be interpreted (and enacted in the real world) because the sequence of linguistic symbols corresponds to real-world sequences.

(5) Take that book and bring it to me.

In this study, I apply the idea of word-order (or constituent-order) iconicity to concessive clauses. It is hypothesized that, if the internal semantic structure of a concessive construction corresponds to a sequence of 'given → new' or 'cause → effect', the surface form of the construction will tend to be arranged in parallel. This applies particularly to the anticausal and epistemic types (see 2.1 above). For example, in (6) there is no strictly temporal ordering of propositional content, but a conditional one: 'If you are short, you are more likely to be overlooked'. From a different perspective we could say that there is a more general, given background condition ('Janet is shorter than Bill') which is expected to take effect in real-world situations. Example (6a) would therefore conform to the iconicity principles, while (6b) runs counter to it.

(6) a. Although Janet was shorter than Bill, she was always noticed first.

b. Janet was always noticed first, although she was shorter than Bill.

Simone (1995, pp. 164–165) also discusses relations of conditionality in these terms, arguing that the protasis ('if') is likely to precede the apodosis ('then') in syntax and in the world (Greenberg, 1966, p. 84). In the context of concessives, the iconic arrangement of propositions is expected if this makes their surface form correspond to mental arrangements of conditionality, or cause and effect. In these cases, the subordinate clause would be expected to precede the matrix clause.⁷ If we accept that concessives are cognitively particularly complex (as argued by Kortmann, 1996, pp. 167–175; Di Meola, 1998, p. 348; König, 2006, p. 821), an iconic arrangement of the sentence can be a valuable aid in its decoding.

3.3 | Varieties of English

The model of English that is most important at the analytic (that is quantitative) level is Kachru's (1985, 1988) Three Circles Model. My argumentation, however, will also draw upon Schneider's (2003, 2007) Dynamic Model of Postcolonial Englishes. Since they are both well-known and discussed in uncountable publications, they will be summarized only briefly before the focus is shifted to the predictions I derive from them. The Inner Circle of Englishes comprises varieties which Kachru calls 'the traditional bases of English' where it is 'the primary language', or L1 (Kachru, 1985, p. 12). The Outer Circle comprises varieties of English that have emerged through colonisation, called New Englishes by Platt, Weber, and Ho (1984, pp. 3–4). Here, English is a non-native second language (L2) with some institutionalised role in the speech community (Kachru, 1985, pp. 12–13).⁸ The circles also differ in the way they adhere to norms or standards (Kachru, 1985, pp. 16–17). Inner Circle varieties are norm-providing (or endonormative), while varieties in the Outer Circle are categorised as norm-developing by Kachru, that is they are outward- and inward-looking for their norms. Rather than sorting varieties of English into categories according to status or colonial history, Schneider's (2003, 2007, pp. 21–70) *Dynamic model of the evolution of postcolonial Englishes* (the Dynamic Model, for short) assumes five consecutive developmental phases through which a postcolonial variety may pass. They are: (i) foundation; (ii) exonormative stabilization; (iii) nativization; (iv) endonormative stabilization; and (v) differentiation. In the present

paper, BrE, CanE, and NZE are of course Phase 5 varieties; NigE is classified as Phase 3, while both IndE and PhiE are positioned between Phases 3 and 4 (Schneider, 2007, ch. 5; Schneider, 2011). The progress of varieties through Schneider's developmental stages is inextricably linked to changes in the political and sociolinguistic landscapes of the respective countries: natives depend less and less upon colonisers and move towards political, attitudinal, and linguistic independence. At the level of language, this is reflected in the stabilization of local features (Phase 3), the development of local norms of usage that are (perceived as) relatively homogeneous (Phase 4), and ultimately the development of sociolinguistic continua that are native to the respective territories (Phase 5).

I would argue that in the context of the present study certain limitations are attached to the usefulness of models of English like Kachru's and Schneider's as predictors of variation. In particular, the idea of 'norms', important in both models, needs to be questioned, since there are no explicit (or even implicit) notions of 'correctness' or 'appropriateness' for any of the syntactic or semantic phenomena that are discussed. The syntactic behaviour of subordinate concessive clauses, whether or not conditioned by different semantic types, simply is not captured by categories such as 'grammaticality', '(non-)standardness', or 'stylistic markedness'. For the time being, therefore, I regard the features under investigation as relatively poor variety markers. What is important in my hypotheses, however, is the difference between L1 and L2 varieties in terms of domains of use and rates of use. In L2 varieties, English is used alongside (an)other language(s); it does have an official and institutionalised status, but it is very often not used in all domains, particularly not in private and informal ones – at the very least, there will be strong competition from native languages. In L1 varieties, on the other hand, English clearly dominates in all areas of use, and it will tend to be acquired both in the school and in private, everyday situations. This aspect dovetails nicely with the orientation of the present study: the cognitive factors sketched above (production- and processing-based constraints) are expected to take effect in all varieties; however, they are expected to be less rigid in L1 varieties, whose users are exposed to more exclusively English-based discourse on a daily basis. Those language users, it is argued, will be somewhat better able to produce and process cognitively 'non-ideal' syntactic arrangements, since they have developed the necessary routines through higher rates of use, in more diverse contexts. The prediction would therefore be that, while the general constraints (production, processing, and iconicity) should operate across the board, and the corresponding effects in the model should have the same direction, effects are expected to be weaker in L1 varieties, and thus positional variation is expected to be greater. However, from a usage- or exemplar-based perspective (Bybee, 2006, 2010) one could argue that universal patterns will be strengthened even further if daily exposure to English is greater (in L1 varieties). Accordingly, the prediction developed in this paragraph (formulated as hypothesis 4 in the next section) is the least confident one in this study.

3.4 | Hypotheses

Based on the discussions in this section, my hypotheses for the quantitative analysis are as follows:

1. There will be a general tendency for subordinate clauses to be in final position in the sentence, in accordance with Hawkins's (1990, 1994) 'performance theory of order and constituency'.
2. In speech, final placement is more likely than in writing.
3. Anticausal concessives are more likely to be in non-final position, as this formal arrangement corresponds to the mental arrangement of propositional content. Dialogic concessives will be free to vary or conform to the general tendency of being in final position.
4. Hypotheses 1–3 will apply across the board, but to a somewhat lesser extent in L1 varieties of English.

That is, hypothesis 4 is the only one that will be tested not in terms of the direction of an effect, but in terms of effect sizes. It is also the least confident hypothesis, as explained in the previous section.

TABLE 1 Token numbers of *although* in six varieties of English

	Variety					
	BrE	CanE	NZE	NigE	IndE	PhiE
Spoken	123	57	72	36	27	98
Written	210	162	223	69	65	117
Total	333	219	295	105	92	215

TABLE 2 Models and model comparison using the LOOIC

Model:	0	1a	1b	2	3
LOOIC:	1655.24	1645.51	1608.54	1596.71	1598.64

4 | DATA AND METHOD

4.1 | Corpora, data retrieval, and annotation

Using the concordancing software AntConc (Anthony, 2018), data were retrieved from the six relevant national components of the *International Corpus of English* (ICE; Greenbaum, 1996; Nelson, 1996) by simply searching for the term 'although'. Opaque or fragmented cases were excluded. Observations were coded for semantic type and syntactic structure. The total number of valid cases was $n_{\text{total}} = 1,288$; the full cross-tabulation by variety, mode of production, semantic type, and syntactic position is given in Appendix A. However, it was decided to exclude epistemic concessives due to their rarity ($n = 29$). Thus, the total number of occurrences used in the analyses was $n = 1,259$; their distribution across varieties and mode of production is shown in Table 1. All analyses were conducted in R (R Development Core Team, 2018) and RStudio (RStudio Team, 2009–2018), including the statistical analysis described below. Plotting was done using the R-packages *lattice* (Sarkar, 2008, 2014) and *latticeExtra* (Sarkar & Andrews, 2016).

4.2 | Data analysis

All analyses in Section 5, as well as Figure 1b, are based on Bayesian mixed-effects logistic regression (Jackman, 2009; Kruschke, 2015; McElreath, 2016). Models were fitted using utilities in the R-package *brms* (Bürkner, 2018), which in turn uses Stan (Carpenter et al., 2017; Stan Development Team 2018). The Bayesian approach was preferred over a standard mixed-effects model (Gelman & Hill, 2007; Luke, 2004) as it is not based on specific assumptions concerning sampling distributions, can be informed by previous research, and permits pooling (adjustment) against central tendencies, which results in somewhat more conservative estimates (Kruschke, 2015, pp. 722–725). The outcome variable is FINAL (with the reference category 'non-final'), which comprises initial and medial occurrences of subordinate clauses.⁹ Binary predictors in the main analysis (see Section 5) are centred on zero. They are SPOKEN.CT ('written' = -0.5 ; 'spoken' = $+0.5$) and DIALOGIC.CT ('anticausal' = -0.5 ; 'dialogic' = $+0.5$), as well as their interaction, which takes predictor values of -0.25 (for 'written dialogic' and 'spoken anticausal') and $+0.25$ (for 'written anticausal' and 'spoken dialogic'). Random effects are VARIETY ($n = 6$; see 4.1) and TEXT ($n = 788$). A summary of all variables can be found in Appendix B. Based on the empirical evidence in the literature (Altenberg, 1986, Schützler 2019), moderately informative priors were specified for the effects of SPOKEN.CT and DIALOGIC.CT; for both predictors, previous research suggests favourable effects on the final placement of clauses. Full model specifications for the final model, including priors, can be found in Appendix C. Five models were run and compared, using the leave-one-out information criterion (LOOIC) as specified in Table 2. The alternative models are shown in Appendix E.

Model 3 contains the two main-effect predictors and their interaction, all three of which are permitted to vary randomly across VARIETY. Additionally, DIALOGIC.CT is specified with slopes that vary randomly across TEXT. Despite

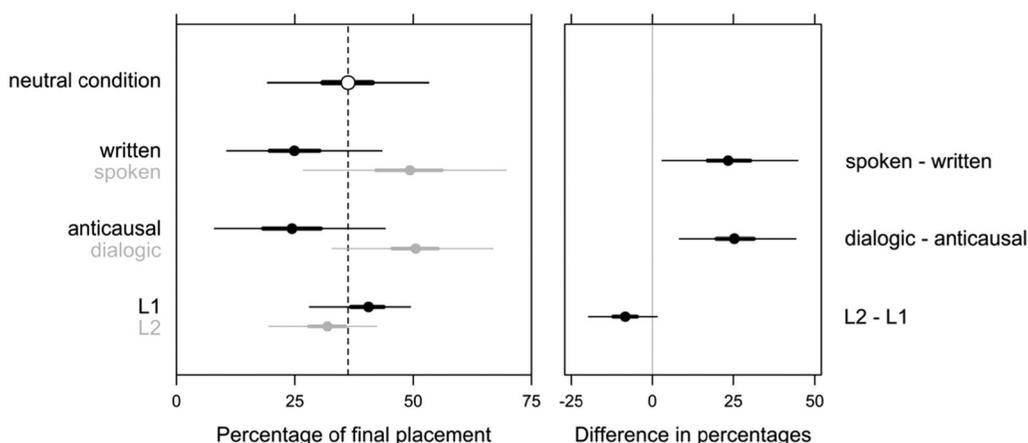


FIGURE 3 Clause positions: Basic conditions based on Model 3 (medians with 50% and 95% uncertainty intervals)

the fact that Model 2 has a slightly lower LOOIC, the maximal model was selected (Barr, Levy, Scheepers, & Tily, 2013).

5 | PREDICTING THE POSITION OF CLAUSES

In a first step, the contrasts in three pairs of conditions will be discussed: (i) writing vs. speech; (ii) anticausal vs. dialogic concessives; and (iii) L1 vs. L2 varieties.¹⁰ The predicted proportions of subordinate clauses in final position are shown in Figure 3, which also contains what is called the 'neutral condition', in which all predictors (with the exception of the global intercept) have values of zero.¹¹ The second panel on the right shows estimates of the three effects (spoken vs. written, dialogic vs. anticausal, and L2 vs. L1). In both parts of the plot, uncertainty intervals indicate how reliable the results are based on the data. In the neutral condition, 36 per cent of concessive clauses with *although* are expected in final position. That is, if we control for mode of production, semantics and variety status, the majority of subordinate clauses will still be in initial (or medial) position. This can be interpreted as first evidence against hypothesis 1 (see 3.4), whereby a general preference of final position was expected on processing-based grounds. There is an astonishingly close match between these findings and Altenberg's (1986) results: if we average across the proportions in LOB and LLC (see Figure 3), the proportion of clauses with *although* in final position is also 36 per cent.¹²

From the global perspective given in Figure 3, the patterns observed for mode of production as well as the semantics of a concessive construction are consistent with hypotheses 2 and 3: in speech, the predicted percentage of subordinate clauses in final position is considerably higher than in writing ($\Delta\% = 23$, 95% UI = [3; 45]).¹³ It has to be said, however, that there is considerable uncertainty in the predicted value in speech, which is indicated by the very long uncertainty intervals in the left-hand part of Figure 3. Nevertheless, the effect sizes shown in the right-hand part of the figure are quite robust, as the uncertainty intervals do not include zero. The difference between anticausal and dialogic concessives is on a very similar scale ($\Delta\% = 25$, 95% UI = [8; 44]). Of course, we still have not considered the interaction between mode of production and semantics, but the results nevertheless suggest clear individual effects of both predictors. The difference between L1 and L2 varieties is not only much smaller, but the comparison also suggests that hypothesis 4 does not hold true on a general level: controlling for semantics and mode of production, subordinate clauses in L1 varieties are somewhat more likely to follow the matrix clause than in L2 varieties ($\Delta\% = 8$; 95% UI = [-2; 20]). This suggests that, if anything, the L1 varieties are affected more strongly by general, processing-based constraints. However, in addition to the unexpected direction of the effect, the 95 per cent uncertainty interval for the effect size includes zero, and we must therefore be particularly careful not to attach too much weight to the

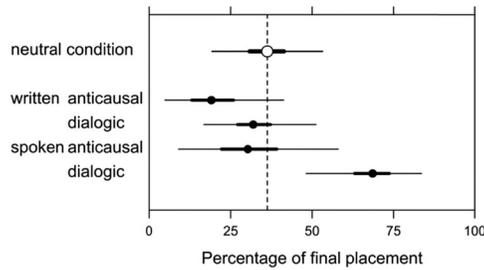


FIGURE 4 Clause positions by mode of production and semantic type (medians with 50% and 95% uncertainty intervals)

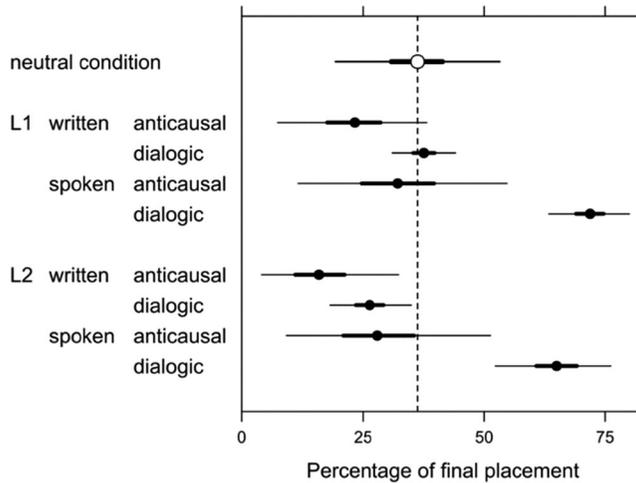


FIGURE 5 Clause positions by variety, mode of production, and semantic type (medians with 50% and 95% uncertainty intervals)

difference between L1 and L2 varieties. These findings call for a reconsideration (or perhaps a general rejection) of hypothesis 4 in future research. Still at a very global level, Figure 4 inspects the interaction of semantic type and mode of production, disregarding the distinction between different types of varieties. As a point of reference, the same neutral condition as in Figure 3 is shown. Scenarios are arranged in a hierarchical order on the vertical axis, with written language on top and spoken language at the bottom, each subdivided into anticausal and dialogic types of concessives.

There is apparently a very marked interaction between mode of production and semantics. The anticausal type of concessive is relatively immune to the effect of spoken language on clause placement: in speech, this type is only somewhat more likely to be realised with subordinate clauses in post-matrix position. The situation is markedly different with regard to dialogic concessives, which are much more likely to have subordinate clauses in final position in speech. Concerning hypotheses 2 and 3 (see 3.4), we therefore need to make the following qualifications: in the spoken mode, the probability of subordinate clauses in final position is indeed higher; however, anticausal semantics act as some kind of brake on this effect. Semantic and production-based constraints are highly interdependent, and, in consequence, dialogic concessives can vary the sequential arrangement of clauses, much more freely, depending on the context of production. Figure 5 takes the analysis one step further still, including the level of variety status. This display does not reveal much in addition to the discussions above, as general patterns are relatively similar in L1 and L2 varieties: we observe the same interaction between semantics and mode of production, the latter taking a much reduced effect for anticausal concessives, and we observe a somewhat increased general probability of final placement in L1 varieties

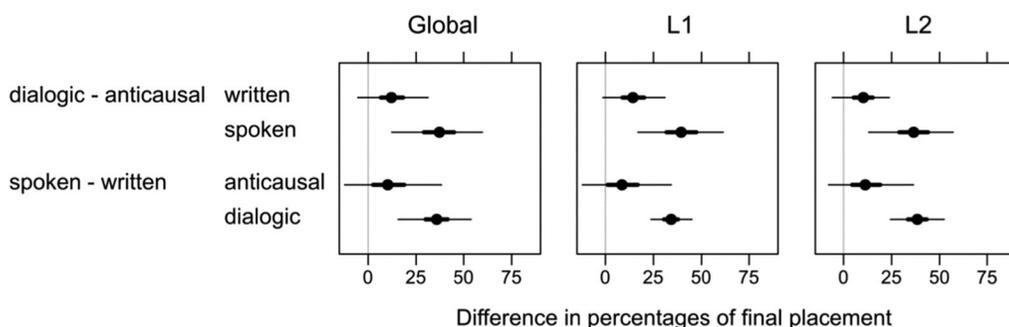


FIGURE 6 Estimates for effect sizes of 'dialogic' (vs. 'anticausal') and 'spoken' (vs. 'written'), shown in interaction (50% and 95% uncertainty intervals)

(also visible in Figure 3 above). That is, semantic and production-based constraints operate in similar ways across all varieties – the difference between L1 and L2 varieties appears to be the weakest one of the tested predictors.

Figure 6 shows estimates for the effects of semantics and mode of production, in interaction with each other, comparing the 'global' condition (disregarding the type of variety) with the situation in L1 and L2 varieties. The first two tiers in each panel answer the question: by how much does the percentage of subordinate clauses in final position change when we move from concessives of the dialogic type to concessives of the anticausal type in writing and in speech? The bottom two tiers show the complementary scenario: by how much does the expected proportion of subordinate clauses in final position change if we move from writing to speech within the subsets of anticausal and dialogic concessives, respectively? Differences between the global, the L1, and the L2 scenarios are very slight and will not feature in the following discussions.

From an alternative perspective, the display illustrates the same findings as Figure 5 above. In addition, it also highlights an interpretational dilemma, namely the problem of causation. In written language, there is only a minor effect of semantics on the position of clauses, and conversely, if anticausal concessives are used in speech and writing, the arrangement of clauses also tends to be affected only little – the respective uncertainty intervals in Figure 6 all include zero. This can mean that written language is immune to semantic effects on clause positions, perhaps due to the retrievability of content in written media: if a text can be processed a second time (by reading it again), clause sequences that are iconic of mental 'if-then' sequences may be less important. On the other hand, the findings may suggest that anticausal concessives are immune to the effects of spoken language, because the iconicity effect is so strong. This may be a genuine case of interacting factors, one supporting the other; it also seems to be the case, however, that the prohibitive effect of anticausal semantics on the influence of speech is somewhat greater than the prohibitive effect of written language on the influence of semantics. The issue can to some extent be resolved if we think of mode of production as primary and semantic function as secondary: when a speaker needs to encode an anticausal or a dialogic concessive, he or she is already producing language that is either spoken or written. Thus, mode (speech vs. writing) is a global parameter that describes the situation in which language is produced. On the other hand, it is much less realistic to think of semantic structure as primary and speech situation/mode of production as secondary. This kind of argument strengthens the interpretation that it is the mode of production that modifies the effect of semantic structure, not vice versa.

Figure 7 is an assessment of the main effects for *SPOKEN.CT* and *DIALOGIC.CT* in individual L1 and L2 varieties. This is an expansion of the information contained in Figures 3 and 4 above. The figure shows by how much (in percentage points) the relative frequency of subordinate clauses in final position is expected to change if conditions are switched from 'anticausal' to 'dialogic', 'written' to 'spoken', and between the two extreme cases, that is from 'written anticausal' to 'spoken dialogic'. Concerning the effect of semantics (Figure 7a), varieties do not differ very markedly, as shown by the parallel lines ('slopes') that link the two conditions. However, they have different baseline values – for example,

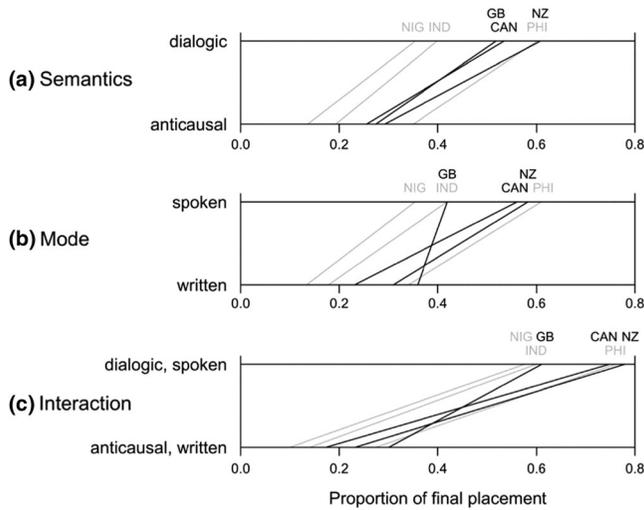


FIGURE 7 Line plots with isolated effects in individual varieties

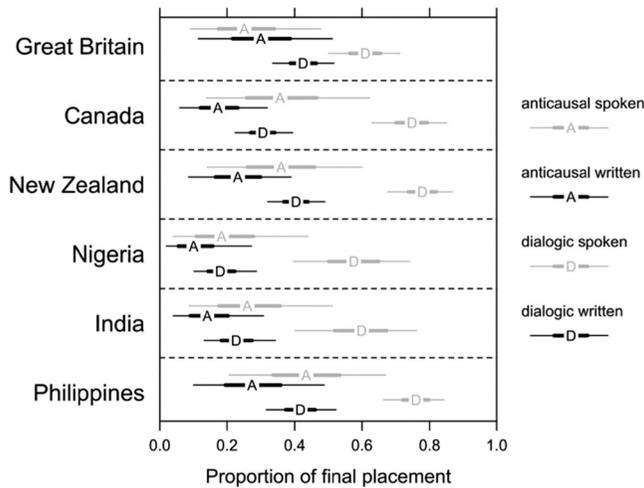


FIGURE 8 Clause positions by variety, semantic type and mode of production (medians with 50% and 89% uncertainty intervals)

NigE and IndE are characterised by rates of final placement that are lower overall. This arrangement, which is also reflected in the other two panels of the figure, indicates that these two varieties are responsible for the lower rate of final placement in L2 varieties evident in Figure 3 above. By and large, the impact of mode of production on clause placement (Figure 7b) is similarly uniform across varieties, with the striking exception of BrE. Here, the effect is still in the same direction (spoken language favouring final placement), but it is considerably weaker. Finally, the comparison of extreme conditions in Figure 7c confirms the similarity of varieties, again with the exception of BrE. In sum, the constraints represented by the predictors *SPOKEN.CT* and *DIALOGIC.CT* are relatively uniform across varieties of the language.

The complete picture is shown in Figure 8, which includes differences between individual varieties, speech and writing, and the two semantic types. The three main findings that were isolated in the analyses above are also evident here: (i) within the spoken and written modes of production, there is always the same tendency for dialogic concessives to be

more likely to occur in final position; (ii) within each semantic category, there is a general tendency for spoken language to produce more final placements, with the notable exception of BrE; and (iii) the spoken-written difference is considerably greater in combination with dialogic concessives. Note that the general impression of great uniformity across varieties is at least partly due to the shrinkage of random effects towards the mean, that is the inbuilt tendency of the model to 'correct' extreme values in the direction of the grand mean.

6 | CONCLUSION

This article investigated the interplay of three different constraints on the placement of subordinate clauses introduced by the conjunction *although* in six L1 and L2 varieties of English. Expectations were formulated as four hypotheses, the first three of which are based on general assumptions that are independent of specific varieties or classes of varieties: (i) holding all other factors constant, subordinate clauses should follow the matrix clause, since this is the ideal arrangement according to Hawkins's (1990, 1994) Performance Theory of Order and Constituency; (ii) due to the greater time constraints placed on advance planning in the spoken mode of language production, subordinate clauses are expected to follow the matrix clause in speech; and (iii) in anticausal concessives characterised by an if-then (or cause-and-effect) relationship between propositions, the syntactic arrangement of clauses will be iconic of that relationship, with the subordinate clause preceding the matrix clause, while in dialogic concessives the subordinate clause will be more likely to post-qualify the matrix clause. Beyond the general interest in these constraints, it was hypothesised that (iv) language use in L1 varieties will be less constrained by the three factors.

A relatively homogeneous picture emerged in the analyses. Across all six varieties, spoken language and dialogic concessives increase the probability of subordinate clauses to be in final position. However, there is an equally uniform interaction effect, whereby written language acts as a brake on the semantically motivated variation of clause positions. In consequence, subordinate clauses are most likely by far to appear in final position if they are produced in speech and are of the dialogic type. Conversely, one could argue that in anticausal concessives the iconicity constraint – resulting in the sequence 'subordinate → matrix' – partly counteracts the natural sequencing of clauses in speech, but I argued above that the first explanatory approach is more plausible. That is, mode of production qualifies the effects of semantics, not vice versa. The fact that this interplay of factors is effective across varieties seems to suggest that it is fairly universal, but, as pointed out earlier, it is probably partly due to the shrinkage of random effects towards the grand mean. Finally, there is no evidence that subordinate clauses are generally preferred in final position. The picture is of course changed if we regard dialogic concessives as the unmarked semantic variant, which could be argued given that this type is more loosely constructed and cognitively less complex, since it does not rest on an underlying if-then relation.

The investigation did not reveal any systematic differences between L1 and L2 varieties. Some varieties (IndE and NigE) are somewhat less likely to produce subordinate clauses in final position, but this is not a large effect, and it does not change any of the other findings discussed above. The fourth hypothesis ('smaller effects for mode of production and semantics in L1 Englishes') was not substantiated, and the constraints that were translated into predictors seem to have some universal (or 'Angloversal'; Mair, 2003) validity. This suggests two conclusions: (i) cognitive, processing- and production-based constraints are strong enough to keep inter-varietal differentiation in check; and (ii) the feature under investigation is not salient enough to develop into a marker of identity. I would propose that the two aspects are likely interrelated: if a linguistic feature is strongly regulated by general cognitive mechanisms, its usage patterns will feel highly natural to speakers and writers of the language, and the emergence of indexical variants will become much less likely. All of these conclusions have to be taken with a grain of salt, however, as systematic effects due to L1 or L2 status are likely to have undergone some degree of shrinkage towards the grand mean. Future research could take several directions, some of which seem more promising than others, however. The investigation of a larger set of varieties (for example using further components of ICE) is of course possible, but the range of varieties included in this

paper is already relatively broad and strongly divergent findings in other varieties would come as a surprise. Corpus studies that target the difference between the two major reference dialects of the language (British and American English) would arguably be more useful. Finally, this kind of study could be expanded to other markers of concession, or, indeed to other types of adverbials.

NOTES

- ¹ This is not to say that correlations between end weight and clause positions are generally ruled out in speech, but it seems more difficult to frame hypotheses based on causal relationships: greater length/weight may well result from (rather than cause) final position in this mode of production.
- ² A subtype of dialogic concessives that could be called 'narrow-scope dialogic concessives' is excluded from the present study. Take, for example, the sentence *We escaped, although only narrowly*. The subordinate structure ('although only narrowly') qualifies not the entire matrix-clause proposition but only the verb phrase. It cannot be re-positioned relative to the matrix clause and thus does not partake in the kind of syntactic variation that is investigated here.
- ³ The display is based on a Bayesian mixed-effects multinomial regression model with mode of production as the only fixed effect, with random slopes for variety and random intercepts for variety and text. The details of this model are not documented here.
- ⁴ The number of tokens in Altenberg's study is quite low, at $n_{LLC} = 20$ and, $n_{LOB} = 49$, respectively. The study is based on a subsample of 100,000 words from each corpus (LLC and LOB).
- ⁵ Also see the cross-linguistic/typological approach taken in Diessel (2001).
- ⁶ Haiman goes on to point out that this view is of course simplistic, since languages may differ very considerably concerning the extent to which they make use of this device.
- ⁷ Interestingly, Altenberg (1984, p. 55) finds no evidence that *because*-clauses follow an iconic arrangement – on the contrary, they clearly prefer final position.
- ⁸ Kachru's Expanding Circle can be disregarded here, as it comprises countries or territories in which a 'performance variety' is spoken, and none of the varieties treated in this study fall into this category.
- ⁹ Clause position was treated as a binary variable since medial positions were rare (Appendix A) and would have put considerable strain both on the model and on the interpretability of results. This decision was also made in previous research (Chafe, 1984, p. 437; Diessel, 2005, p. 452; Wiechmann & Kerz, 2013, pp. 1, 7).
- ¹⁰ For example, to isolate written language, the predictor *SPOKEN.CT* was set to -0.5 , constraining all other predictors to take their (centred) default values, and clause positions were then predicted based on the posterior distribution of coefficients in Model 3. For the contrast between L1 and L2 varieties, predictor values were set to $1/3$ for the random intercepts of the respective varieties in each group.
- ¹¹ This condition represents a (hypothetical) situation in which writing and speech, anticausal and dialogic concessives, as well as L1 and L2 varieties are perfectly balanced.
- ¹² The more precise percentages are 36.3 per cent in my data and 36.4 per cent in Altenberg (1986). Although Altenberg does not inspect the finer details of intra-constructional semantics and does not apply statistical models, the amazingly close replication of his value is reassuring.
- ¹³ ' $\Delta\%$ ' is the percentage point difference between two conditions; '95% UI' denotes the 95 per cent uncertainty interval, that is the least extreme 95 per cent of differences (between the two conditions), based on the posterior distribution of the statistical model.

REFERENCES

- Akinnaso, F. N. (1982). On the differences between spoken and written language. *Language and Speech*, 25(2), 97–125.
- Altenberg, B. (1984). Causal linking in spoken and written English. *Studia Linguistica*, 38, 20–69.
- Altenberg, B. (1986). Contrastive linking in spoken and written English. In G. Tottie & I. Bäcklund (Eds.), *English in speech and writing: A symposium* (pp. 13–40). Stockholm: Almqvist & Wiksell.
- Anscombre, J.-C. (1989). Théorie de l'argumentation, topoï, et structuration discursive. *Revue Québécoise de Linguistique*, 18, 13–55.
- Anthony, L. (2018). AntConc. Version 3.5.7 [computer program]. Available at Retrieved from <http://www.laurenceanthony.net/software.html>
- Azar, M. (1997). Concessive relations as argumentations. *Text*, 17, 301–316.

- Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language*, 68, 255–278.
- Biber, D. (1988). *Variation across speech and writing*. Cambridge: Cambridge University Press.
- Bürkner, P.-C. (2018). *brms: Bayesian Regression Models using Stan*. R package version 2.3.0. Retrieved from <http://cran.r-project.org/web/packages/brms/brms.pdf>
- Bybee, J. L. (2006). From usage to grammar: The mind's response to repetition. *Language*, 82, 711–733.
- Bybee, J. L. (2010). *Language, usage and cognition*. Cambridge: Cambridge University Press.
- Carpenter, B., Gelman, A., Hoffman, M. D., Lee, D., Goodrich, B., Betancourt, M., ... Riddell, A. (2017). Stan: A probabilistic programming language. *Journal of Statistical Software*, 76(1), 1–32.
- Chafe, W. L. (1984). How people use adverbial clauses. *Proceedings of the Tenth Annual Meeting of the Berkeley Linguistics Society*, 437–449.
- Chafe, W. L. (1985). Linguistic differences produced by differences between speaking and writing. In D. R. Olson, N. Torrance, & A. Hildyard (Eds.), *Literacy, language and learning: The nature and consequences of reading and writing* (pp. 105–123). Cambridge: Cambridge University Press.
- Chafe, W. L. (1994). *Discourse, consciousness, and time: The flow and displacement of conscious experience in speaking and writing*. Chicago: University of Chicago Press.
- Chafe, W. L., & Danielewicz, J. (1987). Properties of spoken and written language. In R. Horowitz & S. J. Samuels (Eds.), *Comprehending oral and written language* (pp. 83–113). San Diego, CA: Academic Press.
- Crevels, M. (2000). Concessives on different semantic levels: A typological perspective. In E. Couper-Kuhlen & B. Kortmann (Eds.), *Cause - condition - condition - contrast: Cognitive and discourse perspectives* (pp. 313–339). Berlin: Mouton de Gruyter.
- Crystal, D. (1992). *An encyclopedic dictionary of language and languages*. Oxford: Blackwell.
- Di Meola, C. (1998). Zur Definition einer logisch-semantischen Kategorie: Konzessivität als „versteckte Kausalität“. *Linguistische Berichte*, 175, 329–352.
- Diessel, H. (2001). The ordering distribution of main and adverbial clauses: A typological study. *Language*, 77, 433–455.
- Diessel, H. (2005). Competing motivations for the ordering of main and adverbial clauses. *Linguistics*, 43, 449–470.
- Gelman, A., & Hill, J. (2007). *Data analysis using regression and multilevel/hierarchical models*. Cambridge: Cambridge University Press.
- Greenbaum, S. (1992). A new corpus of English: ICE. In J. Svartvik (Ed.), *Directions in corpus linguistics* (pp. 171–179). Berlin: Mouton de Gruyter.
- Greenbaum, S. (1996). Introducing ICE. In S. Greenbaum (Ed.), *Comparing English worldwide: The International Corpus of English* (pp. 3–12). Oxford: Clarendon Press.
- Greenberg, J. H. (1966). Some universals of language with particular reference to the order of meaningful elements. In J. H. Greenberg (Ed.), *Universals of language* (pp. 73–113). Cambridge, MA: MIT Press.
- Haiman, J. (1985). Introduction. In J. Haiman (Ed.), *Iconicity in syntax* (pp. 1–7). Amsterdam: John Benjamins.
- Haiman, J. (2003). Functional linguistics: Iconicity. In W. J. Frawley (Ed.), *International encyclopedia of linguistics* (Vol. 2, pp. 57–62). Oxford: Oxford University Press.
- Hawkins, J. A. (1990). A parsing theory of word order universals. *Linguistic Inquiry*, 21, 223–261.
- Hawkins, J. A. (1994). *A performance theory of order and constituency*. Cambridge: Cambridge University Press.
- Hermodsson, L. (1994). Der Begriff "konzessiv": Terminologie und Analysen. *Studia Neophilologia*, 66, 59–75.
- Hilpert, M. (2013). *Constructional change in English: Developments in allomorphy, word formation, and syntax*. Cambridge: Cambridge University Press.
- Huddleston, R. D., & Pullum, G. K. (2002). *The Cambridge grammar of the English language*. Cambridge: Cambridge University Press.
- Jackman, S. (2009). *Bayesian analysis for the social sciences*. Chichester: Wiley.
- Johansson, S., Leech, G. N., & Goodluck, H. (1978). *Manual of information to accompany The Lancaster-Oslo/Bergen Corpus of British English, for use with digital computers*. Oslo: University Of Oslo, Department Of English. Retrieved from <http://clu.uni.no/icame/manuals/LOB/INDEX.HTM>
- Kachru, B. B. (1985). Standards, codification and sociolinguistic realism: The English language in the Outer Circle. In R. Quirk & H. G. Widdowson (Eds.), *English in the world: Teaching and learning the language and literatures* (pp. 11–30). Cambridge: Cambridge University Press.
- Kachru, B. B. (1988). The sacred cows of English. *English Today*, 4, 3–8.
- König, E. (1988). Concessive connectives and concessive sentences: Cross-linguistic regularities and pragmatic principles. In J. A. Hawkins (Ed.), *Explaining language universals* (pp. 145–166). Oxford: Blackwell.
- König, E. (1991). Concessive relations as the dual of causal relations. In D. Zaefferer (Ed.), *Semantic universals and universal semantics* (pp. 190–209). Berlin/New York: Foris Publications.
- König, E. (2006). Concessive clauses. In K. Brown (Ed.), *Encyclopedia of language and linguistics* (Vol. 2, pp. 820–824). Amsterdam: Elsevier.

- Kortmann, B. (1996). *Adverbial subordination: A typology and history of adverbial subordinators based on European languages*. Berlin: Mouton de Gruyter.
- Kruschke, J. K. (2015). *Doing Bayesian data analysis: A tutorial with R, JAGS and Stan*. Amsterdam/Boston: Academic Press.
- Linell, P. (2005). *The written language bias in linguistics: Its nature, origins and transformations*. London/New York: Routledge.
- Luke, D. A. (2004). *Multilevel modeling*. Thousand Oaks, CA: Sage Publications.
- Mair, C. (2003). Kreolismen und verbales Identitätsmanagement im geschriebenen jamaikanischen Englisch. In E. Vogel, A. Napp, & W. Lutterer (Eds.), *Zwischen Ausgrenzung und Hybridisierung* (pp. 79–96). Würzburg: Ergon.
- McElreath, R. (2016). *Statistical rethinking: A Bayesian course with examples in R and Stan*. Boca Raton, FL: CRC Press.
- Nelson, G. (1996). The design of the corpus. In S. Greenbaum (Ed.), *Comparing English worldwide: The International Corpus of English* (pp. 27–35). Oxford: Clarendon Press.
- Pawley, A., & Syder, F. H. (1983). Natural selection in syntax: Notes on adaptive variation and change in vernacular and literary grammar. *Journal of Pragmatics*, 7, 551–579.
- Platt, J., Weber, H., & Ho, M. L. (1984). *The new Englishes*. London: Routledge and Kegan Paul.
- Quirk, R., Greenbaum, S., Leech, G., & Svartvik, J. (1985). *A comprehensive grammar of the English language*. London: Arnold.
- R Development Core Team. (2018). *R: A language and environment for statistical computing*. Version 3.5.1. Available at Retrieved from <http://www.R-project.org>
- R Studio Team. (2009–2018). *RStudio: Integrated development for R*. Version 1.1.453. Boston, MA: RStudio, Inc. Available at Retrieved from <http://www.rstudio.com/>
- Sarkar, D. (2008). *Lattice: Multivariate data visualization with R*. Berlin: Springer.
- Sarkar, D. (2014). *Lattice graphics ('lattice')*. R-package version 0.20-29. Retrieved from <http://cran.r-project.org/web/packages/lattice/lattice.pdf>
- Sarkar, D., & Andrews, F. (2016). *Extra graphical utilities based on Lattice ('Lattice extra')*. R-package, v. 0.6–28. Retrieved from <https://cran.r-project.org/web/packages/latticeExtra/latticeExtra.pdf>
- Schneider, E. W. (2003). The dynamics of new Englishes: From identity construction to dialect birth. *Language*, 79, 233–281.
- Schneider, E. W. (2007). *Postcolonial English: Varieties around the world*. Cambridge: Cambridge University Press.
- Schneider, E. W. (2011). *English around the world: An introduction*. Cambridge: Cambridge University Press.
- Schützler, O. (2017). A corpus-based study of concessive conjunctions in three L1-varieties of English. In I. Buchstaller & B. Siebenhaar (Eds.), *Language variation – European perspectives VI* (pp. 173–184). Amsterdam/Philadelphia: John Benjamins.
- Schützler, O. (2018). Grammaticalization and information structure: Two perspectives on diachronic changes of *notwithstanding* in written American English. *English Language and Linguistics*, 22, 101–122.
- Schützler, O. (2019). Variation and change at the interface of syntax and semantics: Concessive conjunctions in American English. In C. Claridge & B. Bös (Eds.), *Developments in English historical morpho-syntax* (pp. 247–268). Amsterdam/Philadelphia: John Benjamins.
- Simone, R. (1995). Iconic aspects of syntax: A pragmatic approach. In R. Simone (Ed.), *Iconicity in language* (pp. 153–169). Amsterdam/Philadelphia: John Benjamins.
- Stan Development Team. (2018). *Stan modeling language users guide and reference manual*. Version 2.18.0. Retrieved from <http://mc-stan.org>
- Svartvik, J., & Quirk, R. (Eds.). (1980). *A corpus of English conversation*. Lund: Gleerup.
- Sweetser, E. E. (1990). *From etymology to pragmatics: Metaphorical and cultural aspects of semantic structure*. Cambridge: Cambridge University Press.
- Wiechmann, D., & Kerz, E. (2013). The positioning of concessive adverbial clauses in English: Assessing the importance of discourse-pragmatic and processing-based constraints. *English Language and Linguistics*, 17, 1–23.

How to cite this article: Schützler O. *Although-constructions in varieties of English*. *World Englishes*. 2020;39:443–461. <https://doi.org/10.1111/weng.12484>

APPENDIX

A Full crosstabs: Occurrences by variety, semantic type and clause position, including epistemic concessives and medial positions

Variety	mode	Anticausal			Epistemic			Dialogic		
		init.	med.	fin.	init.	med.	fin.	init.	med.	fin.
BrE	spoken	11	1	3	4		1	47		61
	written	10	1	11	2			96	4	88
	total	21	2	14	6		1	143	4	149
CanE	spoken	3		3	1			15		36
	written	18	2	6	2			84	7	45
	total	21	2	9	3			99	7	81
NZE	spoken	6	1	3	1		4	16	2	44
	written	20	3	11	1		3	103	2	84
	total	26	4	14	2		7	119	4	128
NigE	spoken	5			2			14		17
	written	4			1		1	49	1	15
	total	9			3		1	63	1	32
IndE	spoken	5		4			1	9		9
	written	9	2	3			2	36	1	14
	total	14	2	7			3	45	1	23
PhiE	spoken	6	1	10	2			23		58
	written	7	2	6	1			50	5	47
	total	13	3	16	3			73	5	105

B Specification of all variables

Variable	Status	Predictor values	Comment
FINAL	outcome (Models 0–3)	'non-final' 'final'	reference: 'non-final'
SPOKEN.CT	fixed (Models 1b–3)	'spoken': 0.5 'written': –0.5	centred at zero
DIALOGIC.CT	fixed (Models 1a, 2, 3)	'dialogic': 0.5 'anticausal': –0.5	centred at zero
SPOKEN.CT:DIALOGIC.CT	fixed (Model 3 Models 2–3)	'spoken, dialogic': 0.25 'spoken, anticausal': –0.25 'written, dialogic': –0.25 'written, anticausal': 0.25	centred at zero
TEXT	Random (Models 0–3)	$n = 788$ levels	
VARIETY	random (Models 0–3)	$n = 6$ levels: 'GB', 'CAN', 'NZ', 'NIG', 'IND', 'PHI'	

C Specification of final model (Model 3)

```
brm(type ~ spoken.ct:dialogic.ct + (spoken.ct:dialogic.ct | variety) + (dialogic.ct | text)
  data = atho.long, family = "categorical",
  prior = c(
    set_prior("normal(0, 2.5)", class = "Intercept"),
    set_prior("normal(0, 2)", class = "b"),
    set_prior("normal(1, 2)", class = "b", coef = spoken.ct"),
    set_prior("normal(1, 2)", class = "b", coef = " dialogic.ct"),
    set_prior("normal(0, 2)", class = "sd")
  ),
  iter = 10000, warmup = 1000, chains = 3, cores = 3, control = list(adapt_delta = .98)
)
```

D Regression table for Model 3 (disregarding effects for individual levels of TEXT)

Coefficient	Mean	Quantiles			n_{eff}	R-hat
		.025	.5	.975		
INTERCEPT	-0.59	-1.43	-.56	0.13	6638	1.00
SPOKEN.CT	1.09	0.14	1.07	2.12	19158	1.00
DIALOGIC.CT	1.20	0.36	1.14	2.34	3504	1.00
SPOKEN.CT:DIALOGIC.CT	0.91	-0.55	0.93	2.24	27000	1.00
VARIETY: Varying intercepts						
Standard deviation	0.71	0.15	0.63	1.73	13937	1.00
GB	0.11	-0.64	0.09	0.97		
CAN	0.09	-0.70	0.08	0.93		
NZ	0.33	-0.40	0.30	1.21		
NIG	-.69	-1.86	-0.63	0.12		
IND	-.36	-1.31	-0.34	0.43		
PHI	0.46	-0.24	0.43	1.39		
VARIETY: varying slopes						
SPOKEN.CT						
Standard deviation	0.81	0.10	0.72	2.00	15002	1.00
GB	-0.82	-2.19	-0.76	0.11		
CAN	0.39	-0.58	0.33	1.64		
NZ	0.08	-0.98	0.06	1.18		
NIG	0.19	-1.02	0.15	1.50		
IND	0.13	-1.03	0.09	1.38		
PHI	0.04	-1.03	0.02	1.14		

Coefficient	Mean	Quantiles			n_{eff}	R-hat
		.025	.5	.975		
DIALOGIC.CT						
Standard deviation	0.50	0.02	0.39	1.63	11730	1.00
GB	-0.11	-1.16	-0.04	0.75		
CAN	0.07	-0.79	0.02	1.06		
NZ	0.20	-0.56	0.10	1.30		
NIG	0.14	-0.93	0.03	1.73		
IND	-0.14	-1.31	-0.05	0.70		
PHI	-0.13	-1.24	-0.05	0.70		
SPOKEN.CT:DIALOGIC.CT						
Standard deviation	0.67	0.03	0.54	2.10	20313	1.00
GB	0.07	-1.33	0.01	1.73		
CAN	0.00	-1.37	0.00	1.36		
NZ	0.11	-1.09	0.03	1.60		
NIG	0.20	-1.24	0.06	2.19		
IND	-0.07	-1.57	-0.02	1.28		
PHI	-0.15	-1.67	-0.05	1.06		
TEXT: Varying intercepts						
Standard deviation	1.57	0.87	1.50	2.61	2362	1.00
TEXT: Varying slopes						
DIALOGIC.CT						
Standard deviation	1.78	0.10	1.69	4.11	937	1.00

E Model syntax for alternative Models (Models 0, 1a, 1b and 2)

Model 2

$final \sim spoken.ct + dialogic.ct + (spoken.ct + dialogic.ct | variety) + (dialogic.ct | text)$

Model 1b

$final \sim spoken.ct + (spoken.ct | variety) + (1 | text)$

Model 1a

$final \sim dialogic.ct + (dialogic.ct | variety) + (dialogic.ct | text)$

Model 0 ('Null model')

$final \sim (1 | variety) + (1 | text)$