Concessive constructions in varieties of English

Ole Schützler

language science press

Language Variation

Editors: Martijn Wieling, Alexandra D'Arcy

In this series:

1. Côté, Marie-Hélène, Remco Knooihuizen and John Nerbonne (eds.). The future of dialects.

2. Schäfer, Lea. Sprachliche Imitation: Jiddisch in der deutschsprachigen Literatur (18.–20. Jahrhundert).

3. Juskan, Martin. Sound change, priming, salience: Producing and perceiving variation in Liverpool English.

4. Dellert, Johannes. Information-theoretic causal inference of lexical flow.

5. Zimmer, Christian (ed.). German(ic) in language contact: Grammatical and sociolinguistic dynamics.

6. Tahmasebi, Nina, Lars Borin, Adam Jatowt, Yang Xu & Simon Hengchen (eds.). Computational approaches to semantic change.

7. Paulsen, Ingrid. The emergence of American English as a discursive variety: Tracing enregisterment processes in nineteenth-century U.S. newspapers.

8. van Gijn, Rik, Hanna Ruch, Max Wahlström & Anja Hasse (eds.). Language contact: Bridging the gap between individual interactions and areal patterns.

9. Schützler, Ole. Concessive constructions in varieties of English.

Concessive constructions in varieties of English

Ole Schützler

Ole Schützler. 2023. *Concessive constructions in varieties of English* (Language Variation 9). Berlin: Language Science Press.

Freie Universität Berlin

To Steffi & Arved

Contents

Contents

Acknowledgements

Abbreviations

A	*although*
a	anticausal
AD/R	addressee/reader
AdjP	adjective phrase
AdvP	adverb phrase
AmE	American English (i.e. English in the USA)
AmE06	*American English 2006 Corpus* (→ xBrown); texts from 2006
ARCHER	*A Representative Corpus of Historical English Registers*; texts from 1600–1999
AusE	Australian English
BBrown	*Before-Brown Corpus* (→ xBrown); texts from 1931
BE06	*British English 2006 Corpus* (→ xBrown); texts from 2006
BLOB	*Before-LOB Corpus* (→ xBrown); texts from 1931
BNC	*British National Corpus*; texts from the early 1990s
BrE	British English
Brown	*Brown Corpus of Standard American English* (→ xBrown); texts from 1961
CanE	Canadian English
CC/CCs	concessive construction(s)
COHA	*Corpus of Historical American English*; texts from 1810–2009
CxG	Construction Grammar
D	difference
d	dialogic
d*	narrow-scope dialogic
E	*even though*
e	epistemic
EFL	English as a foreign language
f	finite
f	frequency
FLOB	*Freiburg-LOB Corpus of British English* (→ xBrown); texts from 1991
fn	final (clause position)

Frown	*Freiburg-Brown Corpus of American English* ($\rightarrow$ xBrown); texts from 1991
HKE	Hong Kong English
ICE	*International Corpus of English*
IndE	Indian English
IrE	Irish English
JamE	Jamaican English
L1	first language / native language
L2	second language
LLC	*London-Lund Corpus of Spoken English*
LOB	*Lancaster-Oslo/Bergen Corpus*
M	mean
n	nonfinite
n	number of observations
nf	nonfinal (clause position)
NigE	Nigerian English
NP	noun phrase
NZE	New Zealand English
OE	Old English
OED	*Oxford English Dictionary*
p	probability; proportion
PDE	Present-day English
PhilE	Philippine English
pmw	occurrences per 1,000,000 words
PP	preposition phrase
Ro	ratio
S	spoken/speech
SD	standard deviation
SEU	*Survey of English Usage*
SingE	Singapore English
SP/W	speaker/writer
T	*though*
VP	verb phrase
W	written/writing
xBrown	extended Brown family of corpora, 1931–2006

1 Introduction

Studies of concessives have been undertaken from mainly three perspectives, as pointed out by Couper-Kuhlen & Thompson (2000: 382–383). Some of the research focuses on concession as the establishment of a particular syntactic link between clauses (or clause-like structures), using concessive connectives. Secondly, concession may be of interest as a semantic text relation, with a focus on the underlying assumptions and semantic mechanisms. A third approach looks at concession from the perspective of rhetoric, that is, it emphasises the role of concession in spoken interaction (cf. Couper-Kuhlen & Kortmann 2000b: 2, Barth 2000, Barth-Weingarten 2003). Those three perspectives broadly correspond to the fields of syntax, text linguistics and discourse analysis; accordingly, certain phenomena and methodologies will take centre stage, depending on the focus that is selected. For example, the syntax-oriented approach will only consider concessives in which an overt connective grammatically marks the relation between matrix clause and dependent structure, while the discourse-analytical approach is much more interested in conceding moves between discourse participants, which may or may not be supported by typical grammatical markers.

The present study is informed by the first two perspectives above, i.e. it focuses on syntactic constructions and the semantic relations that they express (and thus also on local textual coherence). In this approach, concessives only qualify as objects of investigation if they are attached to certain markers, which are in this case restricted to the three conjunctions *although*, *though* and *even though*. The analysis of semantics and syntax is conducted at the level of the construction, not at the discourse level, and a construction is comprised of (i) a semantico-pragmatic relation that holds between two propositions, (ii) two syntactic units (clausal or clause-like) that correspond to propositions, and (iii) a connective. While corpus queries in this study are essentially form-driven, the analysis goes beyond formal aspects (e.g. the counting of markers) and gives semantic and pragmatic considerations a central place. In contrast to discourse-analytical approaches, however, the concrete objects of study are relatively local, in the sense that they do not extend beyond complete sentences and thus treat addressees or interlocutors and their contributions as no more than abstract givens operating

in the background. This can of course be viewed as a shortcoming, but it was considered a necessary restriction; its implications will be discussed in the relevant contexts.

The analytic, quantitative parts of the study are found in Chapters 7–11 of the volume. In Chapters 7 & 8, two surface characteristics of concessive constructions are described in detail: (i) the text frequencies of conjunctions and (ii) the text frequencies of semantic types. These chapters do not establish any relationships between the different functional and formal facets of concessive constructions, and their main function is to prepare and support the more complex scenarios analysed in the later chapters. Particularly assessing the text frequencies of the three conjunctions offers a perspective known from traditional, first-generation corpus-linguistic research, highlighting rates of occurrence without taking recourse to the underlying factors that motivate them, and thus without describing in detail the partly predictable choices made by language users. What places the study as a whole in a Construction Grammar (CxG) context is the approach taken in Chapters 9–11. Here, it is assumed that concessive constructions are characterised by functional and formal properties that matter in combination and therefore need to be explored together. It is of course a challenge for quantitative research to treat the construction as an indivisible whole, rather than focus on one of its characteristics (e.g. semantic type, clause position): Instead of a single variable (e.g. an alternation) I predict the behaviour of constructions comprised of several variable parameters. In this book, I propose a model of constructional choice that rests on five assumptions, as explained in more detail in §4.1.3; as a cognitive model, it will inform the quantitative analyses and the line of argumentation followed in presenting the results.

1.1 Existing research on concessives in English

Characteristics of concessive adverbials are discussed in several edited volumes, which usually treat a wider range of semantic relations, often including or even focusing on languages other than English (e.g. Kortmann 1996, Rudolph 1996, van der Auwera 1998, Couper-Kuhlen & Kortmann 2000a, Ferraresi 2011). There are also many individual articles and chapters, in those volumes and elsewhere, which discuss the semantics of concessives, their relatedness to and overlap with other types of adverbials, as well as the origin and etymology of concessive connectives (e.g. König 1985, 1988, Hermodsson 1994, Azar 1997, Di Meola 1998, König & Siemund 2000). Fewer publications take a quantitative, usage-based approach to concessives; unsurprisingly, all of them approach the topic using

corpus-linguistic methods (e.g. Altenberg 1986, Aarts 1988, Rissanen 2002, Hoffmann 2005, Berlage 2009, Hilpert 2013a). However, all of these contributions are based on present-day and historical British and American Standard English (hereafter: BrE and AmE), and they focus on selected aspects of variation and change (e.g. variable syntactic structures, semantics, frequencies, the choice of connectives), but not on their interaction (or association) in constructions. Hilpert (2013a) is exceptional in going some way towards the inspection of multiple dimensions of variation (semantic types, syntactic realisations), although he, too, stays within the bounds of AmE. What is still lacking, therefore, is research on concessive adverbials that (i) goes beyond BrE and AmE, (ii) inspects several dimensions of variation based on the same data, and (iii) takes steps in the direction of a more holistic CxG account that considers functional and formal criteria in combination. All three points strongly inform the approach taken in this book, which is, moreover, firmly regression-based and transparent in the sense that uncertainties are communicated along with effect sizes. The result is a complex corpus-linguistic research design that generates important insights but also raises interesting methodological questions that can be of value for future research in a CxG framework.

The approach taken in this book may inform research on other (i.e. non-concessive) types of adverbial linkage, although the semantico-pragmatic mechanisms at work in most of them would require considerable adjustments. However, taking the insights gained in this study as starting points for research on other adverbial constructions may be worthwhile since the quantitative and methodological gap outlined above can more or less be argued to hold for the entire domain of adverbial linkage. This becomes evident, for example, if one looks at the amount of research conducted on aspects of the English verb phrase, either based on the central standard dialects (e.g. contributions in Aarts et al. 2013) or actively engaging with World Englishes (e.g. Hundt & Gut 2012). Similarly, the noun phrase in English has been investigated quantitatively with corpora (e.g. Jucker 1993, Pastor-Gómez 2011, Berlage 2014). In contrast, explicitly quantitative studies of adverbial constructions and linkage are few and far between, although this aspect of grammar is certainly rather central. Lenker (2010: 2) therefore still seems to have a point in saying that connectives (and the constructions they bind together) are an understudied area.

1.2 Concessives as constructions

Adverbial (or "circumstantial", König & Siemund 2000: 341) relations expressed by a concessive construction are grouped among "adverbials of contingency"

by Quirk et al. (1985: 479, 484), together with adverbials of reason, purpose, result and condition. In adverbial constructions of this class, two propositions are shown in relation to each other, one of which depends ("is contingent") upon the other (cf. Burnham 1911: 1, Biber et al. 1999: 779). Sometimes it is suggested that, among adverbials, concessives are particularly complex (e.g. Kortmann 1996: 167–175, Di Meola 1998: 348, Hoffmann 2005: 111, König 2006: 821), which has implications for their historical development, their late emergence in language learning, and their cross-linguistic markedness. The particular intra-constructional complexity of concessives is a consideration when formulating hypotheses and expectations prior to the analyses in this book.

Similarly to König (1991b: 633), the present study uses the term *concessive* to refer to the entire bipartite concessive construction – "bipartite" in the sense that it consists of two syntactic structures with propositional content (in this case: clauses or structures interpretable as reduced clauses) that are in some way connected, usually through overt concessive marking. The terms *connective* and *marker* will be used interchangeably: It is assumed that marking a grammatical structure in order to encourage a concessive reading invariably involves connecting (or linking) two components. The concessive as a whole is characterised by at least four variable parameters (or facets), as proposed by Hilpert (2013a: 176): (i) the semantic relationship that holds between the two component parts; (ii) the syntactic arrangement of components (e.g. initial, final or medial placement of the dependent structure relative to the matrix clause); (iii) the selection of the concessive marker(s); and (iv) the internal syntactic form of the subordinate part (e.g. finite clauses vs reduced or nonfinite clauses).[1] Whenever reference is made to a concessive construction (or CC, for short) in this study, it is implied that this entity can be described in terms of these aspects. Crucially, it is assumed that the four dimensions are not independent but linked in a certain way; this view, detailed in §4.1.3, will strongly inform the quantitative analyses in Chapters 9–11 and their sequential arrangement. Some priority is given to functional aspects, which then have formal consequences: It is the primary need to express a certain semantic relation that triggers the selection of a basic syntactic frame (i.e. an arrangement of super- and subordinate), a certain concessive marker and its (finite or nonfinite) clausal complement. I will thus argue that the emergence and entrenchment – and therefore the patterns of use – of concessive constructions can be conceptualised as following a cognitively motivated trajectory, which proceeds from the need to express semantico-pragmatic

---

[1]Hilpert (2013a) focuses on concessive parentheticals, but his four dimensions of variation are undoubtedly applicable to concessives in general.

relations to the formal realisation of these relations, at increasingly fine levels of detail. While formal realisation is of course instantaneous, and therefore happens simultaneously at different levels (clause arrangement, marker selection, realisation of complement), I hope to demonstrate that a cognitively sequential model is helpful, both theoretically and methodologically. Alternative, truly holistic approaches to constructional variation – as foreshadowed in the final chapter – will be explored in future research.

Against the background of this constructional, multifaceted view of CCs, previous research naturally provides an incomplete picture. Thus, Aarts (1988) and Hoffmann (2005) study the frequencies of different concessive markers and analyse their stylistic distribution, but are not concerned with semantic types. Aarts (1988) does investigate the syntactic ordering of CCs marked by certain connectives, but not its interaction with other factors. Berlage (2009), on the other hand, correlates the use of *notwithstanding* as a pre- or postposition with the complexity of the attached noun phrase (NP), but does not differentiate between semantic types either. Hilpert (2013a) includes all four dimensions discussed above in his analysis of concessive parentheticals, which is very much consistent with the CxG framework and its assumptions of an indissoluble link between form and meaning in a construction (hence their definition as *form-meaning pairings*; cf. Goldberg 2003). Therefore, his study is an important point of reference for the present one, although it differs in methodology.

1.3 Aims, scope and structure of the study

Hoffmann (2005: 111) is relatively pessimistic about the feasibility of full-scale studies of concession – that is, onomasiological approaches exploring all possible ways of expressing this relation:

> Given the relatively large range of linguistic realizations, a comprehensive study of concessive relations is certainly not an easy undertaking. This is particularly true given the fact that some sentences may carry a concessive interpretation even though they do not contain an overt marker of concessiveness.

Indeed, the number of possible markers is large, and constructions associated with each of them are potentially characterised by formal and semantic variability across several dimensions (cf. §1.2). Hoffmann also rightly identifies the problem of formally unmarked CCs that are virtually impossible to retrieve automatically from a corpus.

In consequence, the present study does not aim to be comprehensive but, as mentioned above, focuses on the subordinating conjunctions *although*, *though* and *even though*.[2] Concentrating on them as a pseudo-closed set was considered appropriate for the following reasons:

- Unlike conjuncts (e.g. *however*, *nevertheless*), which are always attached to the consequent proposition and follow the antecedent proposition, the conjunctions under investigation partake in the full range of syntactic variation. That is, they (along with their complements) can be in initial, medial and final position relative to the corresponding (in this case: superordinate) component of the CC.

- Unlike other conjunctions that may convey concessive meaning (e.g. *while*, *when*, *if*), they are captured rather well by Sweetser's (1990) semantic framework (cf. §2.2) and do not overlap with other, "primary" (e.g. conditional or temporal) adverbial relations to the same extent (see §2.1.1).

- Due to their etymological relatedness and morphological similarity, they are often treated as a set of quasi-synonymous items; this set, however, is internally underspecified, and its inspection can therefore close a gap in this niche of English grammar.

The general research questions that inform the analyses in this book were partly discussed in §1.2 above. They will be given more substance by formulating hypotheses and expectations in §5.3, which in turn will be addressed empirically in Chapters 9–11, following the more descriptive approaches of Chapters 7 & 8. The underlying broad questions are the following:

1. Is the variation between the final and nonfinal placement of subordinate clauses systematic, i.e. can we account for it in terms of semantico-pragmatics, principles of production and processing, or variety status? [→ Chapter 9]

2. Is the choice between the three conjunctions systematically conditioned by the semantic relation that holds within the construction, or by the general

---

[2]In §7.1, more markers will be cursorily inspected regarding their frequencies in speech and writing. Further, see Schützler (2018c) for a diachronic study of *notwithstanding*; see also Schützler (2018b). A comprehensive treatment of concession would ideally rely on the study of a single, medium-sized corpus, using the automatic retrieval of overtly marked constructions in combination with the manual identification of cases that do not carry an overt marker. This would essentially require reading the corpus.

syntactic frame, i.e. the arrangement of matrix and subordinate clause relative to each other? Do language-external factors (variety, mode of production) affect the choice of conjunction? In short, can the three conjunctions be regarded as quasi-synonymous at all? [→ Chapter 10]

3. Likewise, is the alternation of finite and nonfinite/reduced subordinate clauses systematically affected by the same semantic or contextual factors? In addition, are some conjunctions more likely to attract nonfinite clauses than others? [→ Chapter 11]

These three broad questions will be linked to concrete expectations (or hypotheses) at the end of Chapter 5, immediately before embarking on the quantitative analyses. Concrete expectations are framed relatively late because they depend on the background provided in Chapters 2–4.

The book as a whole is structured as follows: Chapter 2 discusses (i) the etymology and historical development of the markers under investigation, (ii) the different semantic types of concessives that are assumed to exist, and (iii) the forms of concessives both in terms of the position of dependent structures relative to matrix clauses and the internal syntactic structure of complements. Chapter 3 provides corpus examples to illustrate the semantic (and syntactic) categories relevant in the present study and serves as a qualitative counterpoint to the otherwise strongly quantitative analyses. Chapter 4 sketches the theoretical framework of Construction Grammar (CxG) as well as the two dimensions along which variation is mainly explored in this book, namely mode of production and national varieties of English. Chapter 5 presents short summaries of existing research and formulates more concrete expectations and hypotheses on this basis. Chapter 6 lays out the methodologies that were employed. It includes discussions of (i) the corpus material that was used, (ii) the steps that were followed in retrieving, processing and coding the data, and (iii) the applied methods of statistical analysis and visualisation. Chapters 7 & 8 contain the descriptive analyses discussed above, which focus on the text frequencies of markers and semantic types. Chapters 9–11 inspect factors that have an influence on (i) the placement of clauses, (ii) the selection of markers and (iii) the clause-internal syntax of subordinates. These three chapters (and to some extent also Chapters 7 & 8) are essentially parallel in structure and thus provide accessible, in-depth and easy-to-compare treatments of individual aspects. At the same time, they follow the logic of the model of constructional choice presented in §4.1.3. Finally, Chapter 12 contains a general summary of results, discusses their descriptive, theoretical and methodological implications, and points to avenues of future research.

1.4 Open data

The data used in this monograph are published as Schützler (2021) at the *Tromsø Repository of Language and Linguistics* (TROLLing) and can be retrieved via the identifier https://doi.org/10.18710/1JMFVR. Annotated R scripts used for the analyses and visualisations in this monograph can be retrieved from the *Open Science Framework* at https://osf.io/m4tfc/. This repository will occasionally be referred to as "the online appendix", and it also contains all graphics files from this volume. In combination with the original data published at TROLLing, these materials enable readers to rerun all analyses exactly as in this volume, revisualise the data or integrate them into their own research, adapt the models (e.g. by using different priors, including more interactions, or specifying different random effects) or implement altogether different kinds of models (e.g. of a non-Bayesian type) or statistical tests. While individual data tables, scripts and figures at https://osf.io/m4tfc have their unique, direct links, I do not refer to these in the text for the sake of simplicity. However, the repository is structured so as to support easy navigation through the individual components, and there is a ReadMe file explaining how the different parts interrelate.

2 Concessive clauses: Development, function and form

In this chapter, three aspects are addressed. First, §2.1 discusses the close relationship between concessives and other kinds of adverbial relation and shows some of the paths along which concessives (and concessive markers) have developed. Next, §2.2 is an introduction to the semantic categories relevant in the present study. Finally, §2.3 discusses the possible syntactic realisations and the grammatical characteristics of concessive constructions in Present-day English, both in terms of sentence structure and complement-internal syntax.

Following König & Eisenberg (1984: 322; cf. König 1991b: 632, König & Siemund 2000: 341), I use the terms *connective* and *marker* interchangeably. While the focus of this study is on subordinating conjunctions, two other broad classes of connectives can be identified: prepositions and conjuncts (König & Eisenberg 1984: 322, König 1991b: 632, Hoffmann 2005: 110, König 2006: 821).[1] This diversity of grammatically different concessive markers reflects the fact that concessive relations do not exclusively hold between clauses within a sentence, but may involve other structures, for instance nominalisations, entire sentences or larger discourse chunks.

The majority of examples stem from the literature, the *International Corpus of English* (ICE; see §6.1), Brown-family corpora (see beginning of Chapter 3), the *Corpus of Historical American English* (COHA; Davies 2010) or ARCHER (Yáñez-Bouza 2011), and their provenance is cited accordingly. If no further information is given, examples were constructed by the author.

2.1 Historical background

Although diachronic developments of concessive markers play no central role in this study, this section provides some historical background for the contextual-

---

[1]Conjuncts are sometimes referred to simply as "adverbs" (Aarts 1988: 41), "connective adjuncts" (Huddleston & Pullum 2002: 736), "linking adverbials" (Biber et al. 1999: 850–851) or "conjunctional adverbs" (König 2006: 821, 1991b: 632, Hoffmann 2005: 110). For a study of the conjunct *though*, see Schützler 2020b.

isation of the analyses presented in Chapters 7–11. The first part (§2.1.1) of the discussion focuses on the development of English concessives more generally and is followed by short histories of the relevant individual markers (§2.1.2).

2.1.1 General aspects

According to König (1991a: 190), there are (at least) two classes of adverbial relations: (i) "elementary" or "primary" relations (place, time, manner), which can often be "expressed by monomorphemic, non-anaphoric adverbs" (e.g. *there*, *then*, *fast*) and corresponding simple interrogative pronouns (e.g. *where*, *when*, *how*), and (ii) "logical" relations (e.g. causal, concessive, instrumental, and purposive). Historically, logical relations can emerge from primary ones through "secondary grammaticalisation" (Hilpert 2013a: 167–168; cf. examples in König & Traugott 1988: 113–114), but not vice versa. This corresponds to the typical order in which these expressions are acquired by learners (König 1991a: 190–191). In the history of a language, concessives usually develop relatively late, if they develop at all (König 1985: 1–2, 1988: 151, 1991b: 632, Kortmann 1996: 319, König 2006: 821, Hilpert 2013a: 167–168).

König (2006: 821–822; also König 1991a: 192–195) identifies five types of concessive connectives on historical grounds (cf. also König 1985: 10–11 and König & Eisenberg 1984: 323–325, both of whom do not list type 4):

1. Connectives that grammaticalised out of nouns that carry meanings such as 'obstinacy', 'contempt' and 'spite', e.g. the English prepositions *in spite of* and *despite*, which originally depended on some kind of human agentivity (cf. German *trotz*; French *en dépit de*)

2. Connectives one of whose components originated as a free-choice (or universal) quantifier ("allquantor"), e.g. *albeit*, *however*, *whatever*, *anyway*, or *for all that* (cf. German *wie/was auch immer*)

3. Connectives that developed via secondary grammaticalisation, particularly building on temporal, causal and conditional meanings and often combining with an additive focus particle, e.g. *if*, *even if*, *even though*, *even so* (cf. German *obgleich*, *obwohl*, *wenngleich*, *obschon*)

4. Connectives that derive from constructions originally expressing factuality or emphatic affirmation (König cites the construction *True* p, *but* q; cf. German *Zwar…, doch…*)

5. Connectives developed from expressions that highlight a state of remarkable co-occurrence or coexistence, e.g. *nevertheless*, *still*, *notwithstanding* (cf. German *nichtsdestoweniger*, *gleichwohl*).[2]

Connectives like *nichtsdestotrotz* (German) or *in spite of all* show that mixed etymologies exist, in these cases between categories 1 & 5 and categories 1 & 2, respectively. In addition, it can be argued that there is another class of markers that superficially look like nonfinite verb phrases (VPs) but have (partly) grammaticalised into a connective (e.g. *seeing that*, *considering*, *having said that*). The members of this group belong to the category of "marginal subordinators" (e.g. *supposing*, *provided*, Quirk et al. 1985: 1002–1003, cf. Schützler 2018c).[3]

The developmental path of concessive linkers from markers of primary to markers of logical relations (see above) can still be felt in Present-day English (PDE). For example, Aarts (1988: 40) describes concession as "a fuzzy semantic notion", which shades into the neighbouring semantic domains of condition, time and contrast (cf. Quirk et al. 1985: 1088). The overlap between concession and other kinds of adverbial relations is also discussed in Burnham (1911: 66), with a focus on Old English, and in Couper-Kuhlen & Kortmann (2000b: 2; cf. König & Siemund 2000, Harris 1988). As pointed out by Hilpert (2013a: 168), two sources of secondary grammaticalisation are temporal and conditional markers (also cf. Kortmann 1996: 321, Heine & Kuteva 2002: 93, 292), shown in (1) and (2). As in the other examples throughout the present study, connectives will be highlighted in bold print, with their clausal complements in italics.

(1) a. **If** *you try a little harder*, you will succeed. (conditional)
     b. The film was nice, **if** *perhaps a bit cheesy*. (concessive)

(2) a. **While** *living in London*, she decided to become a human rights activist. (temporal)
     b. **While** *clearly a right-leaning person*, he voted socialist on this occasion. (concessive)

If one accepts the primacy of adverbials of place, time and manner, examples like these are symptoms of a process of grammaticalisation in which certain adverbial connectives have developed additional grammatical (namely concessive)

---

[2]Cf. Di Meola (2004: 293–295), who undertakes a similar classification of concessive connectors in German, using properties of their components as criteria.

[3]Very generously (if, of course, also unorthodoxly) one could argue that certain deadjectival lexical items also convey a degree of concessive meaning. Thus, adverbs like *surprisingly* and *unexpectedly* imply the existence of some underlying presupposition which disagrees with the proposition modified by those adverbs.

functions. The older functions continue to exist, which can then be interpreted as a kind of *divergence* (cf. Hopper 1991: 22, Hopper & Traugott 2003: 124–126), whereby primary and secondary grammatical functions can occur alongside each other. At the same time, the principle of *persistence* (see also Hopper 1991) applies, which means that conditional or temporal meanings and associations remain part of a concessive marker's function even after secondary grammaticalisation has taken place.

Two particular types of concessives illustrate the kinship of conditional and concessive adverbial relations. They are what Quirk et al. (1985: 1099–1102) call "alternative conditional-concessive" and "universal conditional-concessive", respectively. Both are subsumed under the category of "irrelevance conditionals" by König (1991b: 635; cf. König & Eisenberg 1984: 315).[4] Example (3a) shows an alternative conditional-concessive. If one of two logically opposed conditions is met (*It is my turn* vs *It is not my turn*), the proposition in the matrix clause will hold true. The focus in this case is of course on the negative condition, which is why the example can be rephrased as shown in (3b) and (3c).

(3) a. In the mornings I scoured the breakfast pans **whether or not** *it was my turn.* (COHA, 1992, fiction)

   b. In the mornings I scoured the breakfast pans **even if** *it was not my turn.*

   c. This morning I scoured the breakfast pans **although** *it was not my turn.*

Examples (4) and (5) illustrate universal conditional-concessives.[5] This type of concessive occurs in combination with the marker *however*, which thus has two possible functions, either as a special kind of "fused" conjunction (as in these cases) or as a conjunct.[6] In the two examples, there are not two alternatives as in (3), but "any number of choices" (Quirk et al. 1985: 1101), including those that would under normal circumstances prevent what is stated in the matrix clause proposition.

(4) [H]owever *hard she strove*, she could not suppress a slight quivering of her lips. (COHA, 1877, fiction)

---

[4]Thompson et al. (2007: 262–263) use the term "indefinite concessive" instead of Quirk et al.'s (1985) "universal conditional-concessive", because such constructions "contain some unspecified element". Other concessives they call "definite concessive".

[5]Hermodsson (1994: 64–65) calls concessives like German *Was auch geschieht...* "generell-inkonditional", a category he developed in an earlier study (Hermodsson 1978: 80).

[6]I call *however* and *whatever* "fused" in such contexts, because they appear to be conjunctions and components of the following clause at the same time.

(5)   **Whatever** *I say to them*, I can't keep them quiet. (Quirk et al. 1985: 1101)

According to König (1991b: 638), constructions of this kind, as well as conditionals more generally, are one important source construction for present-day concessives.

Causal adverbials are less often mentioned in the context of secondary grammaticalisation (Hilpert 2013a: 168), although the connection between causality and concession is often pointed out (e.g. Verhagen 2000: 373–375). However, example (6) demonstrates that, like conditional or temporal meaning, causal meaning can also develop into concessive meaning.

(6)   a. He couldn't move **for** *fear*. (causal)
      b. She loved him **for all** *his faults*. (concessive)

Although it can be argued that *for all* is a complex connective different from simple *for*, the connection between concession and cause is nevertheless evident. Example (7) illustrates another interesting construction which blends causality and concession.[7]

(7)   **Just because** *the lights are on* doesn't mean that John is in his office.
      (from Hilpert 2007: 31; cf. Hilpert 2013a: 168)

The development of concessive connectives out of other markers is also reflected in typology: Connectives with a truly and uniquely concessive meaning do not seem to exist in all languages, while adversative coordinating conjunctions (German *aber*; English *but*) seem to be very common (König 1991b: 632). It is also possible for a language to rely on the context of an utterance to disambiguate a multifunctional connective. In English, expressions with a clear concessive meaning exist alongside markers in which primary grammatical functions persist. Aarts (1988: 40) calls the former "centrally concessive" (e.g. *although*) and the latter "peripherally concessive" (e.g. *whereas*, *if*).[8]

2.1.2 The histories of *although*, *though* and *even though*

According to the *Oxford English Dictionary* (OED; *s.v.* "though"), Old English (OE) *þéah* – or one of its variants – seems to have been the original form out of which

---

[7]Very similar constructions exist in German.

[8]Note that Di Meola (1998: 343–348) uses the label "peripheral" to refer to certain pragmatic types of concessives (cf. §2.2.3).

the etymologically related other items of the set have developed. With the exception of certain dialects (e.g. East Anglian), this was replaced in Middle English by the form *þóh* (or one of its variants), which derived from Old Norse and had a back vowel. These forms were the basis for developments in the standard language up to PDE. As early as OE, it was possible to add a preceding intensifying form, *eall* (cf. Burnham 1911: 12–14, Eitle 1914: 114, Chen 2000: 104–105). In Middle English, this became *alle*/*all*/*al*, either free-standing or hyphenated to a variant form of *though*. Original and expanded forms have coexisted up to the present day. The development of *although* constitutes a case of grammaticalisation, with the particle *alle*/*all*/*al* not only becoming firmly attached to *though* but also losing its intensifying character. It is interesting to note that the first uses of OE *þéah*/*þéh* as conjuncts (or sentence adverbs) date from roughly the same time as its use as a conjunction (cf. Eitle 1914: 112), and apparently OE did not make a clear syntactic distinction between the two uses. The OED states that, like *although*, *even though* came to be used as an emphatic, intensifying variant.

The OED establishes a relatively clear chronology. The earliest attestations of the predecessors of *though* as a conjunction are given for the 9th and 10th centuries. A variant approximating the modern, Norse-based form (*þou*) is cited from the 14th century. It is around the same time that *although* as a conjunction seems to have arisen, if of course in variant spellings. Finally, *even though* is attested considerably later, in 1697.

In PDE, *although* – a marked (emphatic) form at the time of its emergence (cf. Burnham 1911: 19–20, Bryant 1962: 216) – is the most frequent of the three conjunctions (see, for instance, results in Schützler 2017). While it no longer stands out as an emphatic variant, *although* may have developed a different kind of markedness, if we accept Quirk et al.'s (1985) claim that it is more formal than *though* (see Chapter 5). That is, if *even though* is an emphatic variant (Quirk et al. 1985: 1099), *although* may be a stylistically (slightly) elevated variant.

2.2 Semantic types of concessives

The analyses in this study reckon with three semantic types of concessives, as discussed by Sweetser (1990: 76–78).[9] Sweetser's "content" and "speech-act" types

---

[9]There are some contributions that overlap with (and partly antedate) Sweetser's (1990) influential three-way categorisation. See, for example, Borkin's (1980: 51) distinction between a "dissonance of an empirical nature" and a "dissonance of a rhetorical nature" in concessives, or Halliday & Hasan's (1976: 250–253) discussion of "external" and "internal" adversatives.

will be referred to as *anticausal* and *dialogic*, respectively, while the label "epistemic" is left unchanged. Furthermore, the dialogic type is subdivided into two variants, as follows:

1. anticausal concessives

2. epistemic concessives

3. dialogic concessives:

    a) wide scope

    b) narrow scope[10]

Using these types goes beyond more general definitions as provided by Quirk et al. (1985: 1098), for example, according to whom "[c]oncessive clauses indicate that the situation in the matrix clause is contrary to expectation in the light of what is said in the concessive clause". While all concessives share an element of surprise or unexpectedness, the more fine-grained semantic categories are needed to describe them precisely.

The following four examples from the *International Corpus of English* (ICE; see §6.1) illustrate the four semantic types. The main points of difference will be highlighted with short, non-technical paraphrases, while more theoretical and detailed discussions will be provided in the respective individual sections below.

(8) Anticausal concessive:
     Although *no one will publish his work*, he refuses to compromise his artistic integrity by catering to the marketplace. (ICE-CAN:W2F-019)

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

(10) Dialogic concessive (wide scope):
     On the basis of this material historians have been able to establish some unassailable facts, **although** *they are not agreed on how this evidence should be interpreted.* (ICE-IRE:S2B-036)

(11) Dialogic concessive (narrow scope):
     I have tried to phone you a couple of times this week, **although** *not persistently.* (ICE-GB:W1B-001)

---

[10]These labels and the subdivision of dialogic concessives were introduced in Schützler (2020a). For further semantic categories of adverbial linking, see Crevels (2000: 315–317), Lang (2000) and Tsunoda (2012).

Example (8) is an anticausal concessive, whose semantic structure can be para-phrased as follows: 'He is unsuccessful as a writer, but, *unexpectedly, this does not result in a change of writing style or subject matter.*' The italic part of the sentence points to what will be called a *topos* in §2.2.1 below, i.e. the assump-tion that, under normal circumstances, a certain set of circumstances will have certain consequences or results. The term *anticausal* refers to the fact that the causal trajectory that is triggered is not consonant with the presented facts. That is, preconceptions held by the addressee or reader concerning causes and effects are activated for the decoding of the concessive, but the normally assumed cause-and-effect relation remains unrealised.

The epistemic concessive in (9) can be paraphrased as follows: 'Pictures always show Bonifacio [a Filipino revolutionary leader; OS] with a bolo, but, *although this portrayal will naturally make us think so, he never actually fought with this kind of weapon.*' The dependent clause in (9) expresses observed facts or phenom-ena that suggest or encourage (or, indeed, cause) certain conclusions. However, the observed facts cannot be construed as real-world causes (the brandishing of bolos – presumably in portraits – cannot possibly cause Bonifacio to have used them in the past).

Inferences of the types discussed above (i.e. inferences concerning either causes or effects) are not central in constructing and decoding dialogic conces-sives like (10), which can be paraphrased as 'On the one hand, there is agreement concerning the facts of the matter – this is helpful; on the other hand, historians are not agreed concerning what those facts mean – this complicates things.' In this case, two propositions make differently-angled contributions to the overall evaluation of a situation. In the example, the positive tone of the first one is damp-ened by the second one. Referring to such constructions as *dialogic concessives* highlights the fact that the two propositions enter into some kind of dialogue, in the sense that both are subject to reciprocal pragmatic qualification and modifi-cation. The relationship between propositions might be argued to be adversative, rather than concessive in the strict sense of the word.

Finally, narrow-scope dialogic concessives are regarded not as an entirely in-dependent category but as a subtype of dialogic concessives. Like in (10), the subclause in (11) modifies the proposition in the matrix clause without trigger-ing causal inferences. In this case, however, the (adverb) phrase introduced by *although* does not have scope over the entire main clause but only over its verb phrase. While dialogic in nature, narrow-scope dialogic concessives are treated as a separate category, since – like in (11) – the dependent clause does not present a new proposition, but essentially functions like a negatively-phrased adverbial

of manner. Table 2.1 summarises the main traits of the different semantic types. More detailed discussions will follow in §2.2.1–2.2.4.[11]

Table 2.1: Semantic types of concessives

Type	Short description
anticausal	Proposition in dependent structure presents real-world obstacle to matrix clause proposition; understood and decoded on the basis of a topos, i.e. an assumed "normal" relation of cause and effect
epistemic	Proposition in dependent structure suggests certain conclusions; typically also based on a topos; often functions like an inverted anticausal concessive, i.e. the observed effect suggests an assumed cause
dialogic (wide scope)	Propositions in dependent structure and matrix clause present different perspectives on a single situation; proposition in dependent structure qualifies, modifies or corrects the matrix-clause proposition, but stands in no relation of cause and effect
dialogic (narrow scope)	Semantic/pragmatic function like dialogic concessives; only part of the matrix-clause proposition (e.g. the VP and its complements) is qualified, modified or corrected

Like the syntactic categories discussed in §2.3, semantic categories will be simplified for the central quantitative analyses by including only anticausal and wide-scope dialogic CCs. Epistemic concessives, while theoretically interesting, are rare in the data, and their inclusion in regression models would generate more problems than real insights. Narrow-scope dialogic CCs – apart from also being relatively rare – are syntactically highly restricted, i.e. they hardly partake in formal variation as defined in this study.

---

[11]For a discussion of the different semantic types of concessives regarding their degrees of subjectivity, see Crevels (2000), Hilpert (2013a) and Schützler (2018a); for general discussions of subjectivity that can contribute to this kind of approach, see Benveniste (1971), Halliday & Hasan (1976: 26–27), Traugott (1989, 2010, 2014) and Langacker (1985, 1990).

2.2.1 Anticausal concessives

As already outlined above, an anticausal concessive is constructed and decoded based on a *topos* (Azar 1997: 306, Anscombre 1989), which is a presupposition in the form of an IF → THEN relation shared (i.e. understood) by the speaker or writer (for short: SP/W) and the addressee or reader (AD/R; cf. König 2006, Givón 1990: 835).[12] Topoi can be very general or nearly universal, in which case they require very little (or no) context; on the other hand, they can also be highly context-specific, in which case the topos is valid only for a particular communicative situation, a certain time period, or a certain speech community or culture. An example of a universal topos is perhaps LITTLE SLEEP → TIREDNESS, i.e. 'if you sleep little, you will be tired'.[13] It is reasonable to assume that this mechanism will be operative irrespective of time, place and social factors, because it is part of the physical human condition. On the other hand, a topos may also be more restricted, e.g. regionally or historically. Take, for example, the construction *Although only 22 years old, he was allowed to vote*. This would make little sense in present-day western societies. In late nineteenth-century Prussia, however, men were allowed to vote only if they were at least 24 years old, so different assumptions hold for this particular historical political system, making the above sentence perfectly functional and easy to decode in that context.

The view of concessives as based on causal or conditional relationships is also reflected in Quirk et al.'s (1985: 484) treatment of concession as "an 'inverted' condition" or "a 'blocked' or inoperative cause", as well as in Halliday & Matthiessen's (2004: 272) description of concessives as construing "frustrated cause". The term *anticausal* in the present study is intended to be a more transparent reflection of this underlying mechanism than the term "content concessive" (Sweetser 1990, Crevels 2000, Hilpert 2013a: 78).

Example (12) hinges upon the (general) topos MAKING HASTE → PUNCTUALITY.[14] It appears sensible to describe topoi by using maximally general formulations like this, so that they can capture a large number of actual realisations (cf. König 1991b: 633, Hermodsson 1994: 73).

(12)   **Although** *I ran fast*, I missed the bus.

---

[12]There is a wealth of alternative terms, e.g. "hypothesis" (Burnham 1911: 1–2), "presupposition" (König 1991a), and "assumption" (König 2006).

[13]I will use small caps to highlight generalised relationships between propositions in concessive constructions.

[14]Other typical topoi would be HARD WORK → SUCCESS, LITTLE TO DRINK → THIRST, and ACTIVE SOCIAL LIFE → NOT FEELING LONELY.

Example (13) is reproduced from Sweetser (1990: 79). Someone who is not aware of an emergency (because they have not heard the call for help) will normally not come to the rescue. A more general topos on which the construction is based could be formulated as UNAWARENESS OF PROBLEM → INACTIVITY.

(13)   **Although** *he didn't hear me calling*, he came and saved my life.

In the terminological framework proposed by Azar (1997: 308), anticausal concessives are "persuaders". They do not provide additional evidence in favour of the (unmarked) primary statement but make it more convincing: Firstly, they anticipate, make explicit and thus disarm facts that might otherwise be used to undermine or discredit the main proposition; secondly, they increase the credibility of SP/W, who presents a more multifaceted, balanced and complete picture of the situation and thus comes across as circumspect and thorough. Di Meola (1998: 341) argues that by mentioning an obstacle in the antecedent, the consequent proposition is highlighted, appearing less natural. In addition, AD/R's curiosity may be piqued, and their thoughts may be directed towards a yet unknown cause for the non-realisation of the default causality, thus potentially contributing to the coherence of a text by foreshadowing further evidence provided in subsequent parts of the discourse.

The cause in a topos may be construed indirectly, as in the following example. Here, tipsiness does not result directly in incomprehensible speech. Rather, it may result in slurred or indistinct speech, which in turn is likely to make spoken messages incomprehensible.

(14)   But we gathered, **although** *they were tipsy*, that their first names were Lily and May. (ICE-IRE:S2B-021)

The logical chain of causal relations in the example could run as follows: (i) 'Because they were tipsy, their speech was indistinct'; (ii) 'although their speech was indistinct, we understood that their names were Lily and May'. Thus, two topoi are effectively fused into one: DRUNKENNESS → OBSCURE SPEECH → COMMUNICATION PROBLEMS. The intermediate 'obscure speech' element is left entirely implicit, but SP/W can safely rely on AD/R's ability to fill in the gap, based on their world knowledge.

2.2.2 Epistemic concessives

Like anticausal concessives, epistemic concessives are often based on topoi. In this case, however, the two propositions are not held together by an IF → THEN

relation – at least not in the same way as in an anticausal CC. Hilpert (2013a: 165–167) describes the difference as follows: Epistemic concessives, like anticausal concessives, invoke a "causality frame" (which can, for practical purposes, be equated with a topos; see §2.2.1 above). In epistemic concessives, however, the causality frame is not based on "real-world causation" but on "inference" (Hilpert 2013a: 165; cf. Crevels 2000: 318).

In (15), observing that someone has failed her final exams may lead to certain conclusions, among them perhaps that the person in question is not a particularly gifted student. This conclusion, however, turns out not to be in harmony with reality in this case:

(15)   She is a very clever student, **although** *she failed her finals.*

Two important notes need to be made. First, it is crucial that the conclusion based on the proposition in the subordinate clause is what could be called *regressive*, i.e. it concerns states of affairs (or processes and actions) that are prior to (or underlying) the italicised proposition. In the example, the relevant relationship is not between failing the exams and its possible real-world consequences, but between failing the exams and possible causes, facts or personality traits that can account for it. Secondly – and this characteristic is shared between anticausal and epistemic concessives – we cannot assume that the content of the subclause triggers a highly specific conclusion. Rather, the semantic structure of the construction as a whole is such that the main-clause proposition contrasts with one of numerous possible inferences triggered by the proposition in the subordinate. Thus, in the example, failing one's exams could be due to a lack of talent, preparation, interest in the subject, or physical/mental fitness.

Example (16) is taken from Sweetser (1990: 79), where it is presented along with the sentence reproduced as (13) above. The propositional content is the same, but in (16) the past perfect is used to make explicit that the subordinate clause is shown in relation to a prior event (see comments above).

(16)   **Although** *he came and saved me*, he hadn't heard me calling for help.

As the examples in this and the previous section show, an anticausal concessive can in many cases be turned into an epistemic concessive (and vice versa) by re-attaching the concessive marker and thus changing the status of main and subclause, possibly supported by an additional adjustment of tenses. On the basis of examples like these, epistemic concessives can be regarded as anticausal

concessives with inverted semantic polarity.[15] In the same vein, Di Meola (1998: 345–346) calls epistemic concessives "reconstructive" and points out that they are characterised by a reversal of the "argumentative direction".[16]

2.2.3 Dialogic concessives

In this study, the term *dialogic* will be used for a relatively broad category of concessives. What they have in common is the absence of inferences concerning likely effects and outcomes (as in anticausal CCs) or likely underlying causes or states of affairs (as in epistemic CCs). Instead, in the most general definition of dialogic concessives, one of the two propositions qualifies the other, qualitatively or in degree, or both propositions present conflicting evidence and thus suggest different courses of action or different evaluations of the situation as a whole. Sweetser (1990), Crevels (2000) and Hilpert (2013a) use the term "speech-act concessive" for this type of construction.[17] I would argue that, although this term is a good label for certain types of concessives, a broader designation is needed, particularly since different notions exist as to what precisely constitutes a speech act. Still, the term *dialogic concessive* as I use it is co-extensive in meaning with the term "speech-act concessive", and therefore Hilpert's (2013a: 167) following definition of speech-act concessives clearly applies here:

> With the first element, the speaker makes a pragmatic commitment that would, in a default scenario, cause her or him to make subsequent statements consistent with that commitment. Yet, the commitment is withdrawn, and this is signalled with the use of a concessive conjunction.

The withdrawal of pragmatic commitment described by Hilpert is what I call *qualification* or *modification*: One proposition sets the discourse off on a certain pragmatic trajectory, suggesting certain evaluations or courses of action, while the second proposition qualifies, weakens, or indeed cancels that pragmatic trajectory. Crevels (2000: 318) argues more purely in terms of speech acts:

---

[15]Both Quirk et al. (1985: 1098) and Huddleston & Pullum (2002: 735) point out that the concessive marker can be attached to either of two clauses. However, only Huddleston & Pullum discuss the fact that moving the connective from the head of one clause to the head of another changes "[t]he implicature" – they effectively describe the difference between anticausal and epistemic types, without using those terms.

[16]German: "rekonstruktiv"; "Argumentationsrichtung".

[17]Publishing her book in the year of J. L. Austin's birth, Burnham (1911: 33) naturally does not use the term "speech act", although she describes the speech-act type (which I call *dialogic*) when she writes that OE concessives with *þeah* are "sometimes used loosely to relate or contrast two ideas between which there is no logical opposition", in which case the concessive clause "is added simply as a qualifier".

> In the speech-act domain the content of the concessive clause does not form an obstacle for the realization of the event or the state of affairs described in the main clause, but raises obstacles for the realization of the speech act expressed by the speaker in the main clause.

Any qualification, weakening or withdrawal of "pragmatic commitment" (Hilpert 2013a: 167) can be regarded as an obstacle to the realisation of a speech act, if speech acts are defined broadly enough. As a motivation of the term *dialogic*, it could be argued that the two propositions in constructions of this type enter into a dialogue with each other, one of them promoting certain evaluations or courses of action, the other providing additional and potentially conflicting information with an impact on how the concessive as a whole is to be interpreted. At the level of discourse, it could further be argued that SP/W presents an unresolved situation, and that meaning-making ultimately depends on how AD/R engages with this. Thus, the process involves an inter-propositional dialogue as well as an intersubjective one.

A prototypical case of what is called "speech-act concessive" in the literature is shown in (17): The unmarked clause contains a declaration (*I'm innocent*), possibly meant to encourage a supportive course of action in AD/R; at the same time, however, SP/W presents a second proposition (*I know you won't believe me*) which reduces the probability of the matrix-clause speech act being felicitous. There is no topos-based, factual incompatibility between being innocent and expecting to be disbelieved. The relatedness of different types of CCs becomes once again evident if we expand the matrix clause in the example (*I am saying that I'm innocent, although...*), for instance.

(17)   I'm innocent, **although** *I know you won't believe me*. (from Sweetser 1990: 81)

Example (18) from ICE-GB is another instance that is not only dialogic but truly speech-act in nature. Here, the unmarked clause contains the writer's birthday congratulations, while the concessive clause expresses the certainty that the letter expressing them will arrive late, effectively making the congratulating speech act less felicitous.

(18)   Happy 25th Birthday for Monday, **although** *this letter will arrive days and days after your birthday*. (ICE-GB:W1B-005; comma added)

In (19a), taken from Hilpert (2013a: 165), the speech act is less clearly identifiable as such. It is the frequent occurrence of CCs of this kind that led me to

explore alternative, more general labels for this semantic category, resulting in the term *dialogic* concessive.

(19)   a. **Although** *surgery is best*, it is not always possible. (Hilpert 2013a: 165)
        b. Surgery is best, **although** *it is not always possible.*

The proposition in the subordinate clause of (19a) focuses on the fact that surgery is the best option available in a certain situation, while the matrix clause states that it is not always feasible. There is no topos whereby the best solution can generally be expected to be viable; what the construction as a whole does is qualify the pragmatic stance of one proposition (implicitly recommending/promoting surgery) by introducing another, which indicates certain complications or restrictions. In constructions of this type, it is possible to simply re-attach the concessive marker and thus change the status of clauses with only minor effects on the function of the construction as a whole, as shown in (19b). I would argue that this is because, unlike anticausal and epistemic concessives, dialogic concessives do not involve inferences in the stricter sense, and thus the link between propositions has no particular directionality.

In the next (constructed) example, two characteristics of a person – good looks and low intelligence – are contrasted, using *although*. There is no conflict between the propositions themselves (beauty vs lack of intelligence), i.e. there is no real-world reason to assume that the two should not co-occur. However, the positive stance of the main clause (presenting the subject as physically attractive) is downgraded by the proposition contained in the subordinate clause (presenting the subject as intellectually *un*attractive). In this case, the concessive relation holds between two evaluative stances, and the final position needs to be negotiated dialogically, based on the evidence. This sentence could be paraphrased as follows: 'His good looks make him an exciting companion; but then again his lack of intelligence might make him boring or even embarrassing company.'[18]

(20)   He is really good-looking, **although** *he's not very bright.*

Finally, in (21) the fact that undernourishment in Argentina is at a relatively low rate (presented as an achievement) is qualified by adding that this is only possible due to state support (i.e. the achievement comes at a cost); the subordinate addition clearly makes the matrix-clause message appear less impressive.

---

[18]Concessives like (20) are called "evaluative" (German "evaluativ") by Di Meola (1998: 345), who provides a similar German example. He (1998: 347–348) also discusses "limiting" and "corrective" CCs (German "limitativ", "korrektiv"), which form a continuum, the latter expressing a stronger qualification or correction than the former. König (2006: 823–824) speaks of "rectifying" concessives, in which "the content of the main clause is weakened". He further claims that the (marked) rectifying clause always follows the matrix clause.

(21) In Argentina, [...] only some 8 per cent of the population is under-nourished, **though** *the National Food Programme is now needed to ensure that food is available.* (ICE-GB:W2A-019)

Dialogic concessives, then, can serve a number of purposes that are not mutually exclusive: (i) They can (ostensibly) express a complex situation in a more objective way by providing contrasting perspectives, which may enhance SP/W's standing in the eyes of AD/R, as they will appear more circumspect and considerate; (ii) they enable SP/W to avoid taking a clear stance and thus responsibility for consequent actions and decisions; and (iii) they can give AD/R more interpretative leeway. Because SP/W avoids taking an entirely clear stance and several interpretations are possible, dialogic concessives are pragmatically "mixed messages" (Hilpert 2013a: 166).

2.2.4 Narrow-scope dialogic concessives

As the label suggests, narrow-scope dialogic concessives are treated as a subtype of dialogic concessives. They are more limited in semantic scope and the dependent structure lacks syntactic mobility.

The narrow semantic scope of this type of CC can be seen in (22). The dependent negative adverb phrase introduced by the connective does not comment on the entire matrix clause proposition but constitutes a qualifying addition to one aspect only, namely the degree of improvement.

(22) It improved on a standard Philips design, **though** *not a great deal.*
(ICE-GB:W2B-040)

Example (23) is perhaps an even clearer illustration of this semantic type: *reluctantly* is a modification of the VP (*agreed*). In order to give the dependent part of the CC scope over the entire matrix clause, one could use an adjective phrase (AdjP) instead of an adverb phrase (AdvP): *Though reluctant, the child agreed.* Alternatively, one could restate the subject along with a resumptive (dummy) predicate: *...though she did so reluctantly.*

(23) The child agreed, **though** *reluctantly.* (ICE-IND:W2B-018)

Narrow-scope concessives will be treated as strictly of the dialogic type and strictly nonfinite.[19] Concerning syntactic ordering in narrow-scope CCs, it is almost categorically the case that the connective and its complement – usually an

---

[19]As discussed in the text, it is possible to re-construct (23) using a finite clause (*...though she did so reluctantly*), but this kind of construction simply did not occur in the data.

AdvP or a preposition phrase (PP) – follow the matrix clause. Rearrangements will result in ungrammatical constructions (*Though not very quickly he answered the phone*; *He though not very quickly answered the phone*). In ICE-Philippines – not included in the quantitative part of this study – there was a single example in which the typical sequence of elements was inverted, as shown in (24). The canonical form would be either *significantly, though not fully*, or perhaps *not fully, though significantly*. The fact that this occurs in an L2 variety is in accordance with the finding that (some of) those varieties may treat connectives somewhat differently, sometimes using a second, correlative marker (in this case: *though... but...*).[20]

(24)   This phenomenon was, **though** *not fully* **but** significantly, explained by the Sapir-Whorf theory, […]. (ICE-PHI:W1A-007)

Examples like (25) and (26) were also categorised as narrow-scope, even though they function somewhat differently. In both cases, the attribute of a following noun is postmodified by an AdjP marked for concession.

(25)   [A]nd the study of Latin occupied a subsidiary, **though** *nonetheless important*, place in the curriculum of the Scottish universities. (ICE-CAN:W2A-008)

(26)   They have a surprisingly loud, **though** *squeaky*, voice for so tiny a bird. (ICE-JAM:W2B-021)

Crucially, all examples in this section are characterised by subclausal postmodification, be it within AdvPs, AdjPs or relative to entire VPs. Secondly, rigid constraints are in place concerning the syntactic placement of narrow-scope CCs. And, finally, it is grammatically not possible for narrow-scope dialogic CCs to be constructed with finite clauses, which is a direct result of subclausal status. Rather, CCs of this type employ an AdjP (if postmodifying an AdjP) or an AdvP/PP (if postmodifying an AdvP or a VP).

2.3 The syntax of concessive subordination

This section deals with the syntactic properties of CCs with *although*, *though* and *even though*. Section 2.3.1 discusses the general aspect of syntactic ordering, i.e. the positioning of subclauses relative to matrix clauses, while §2.3.2 focuses

---

[20]See, for instance, (69) and (70) from IndE and HKE (§3.5, p. 50).

on basic types of clauses (or clause-equivalent structures) that combine with the three connectives.

Subordinating conjunctions introduce clauses that depend on another clause, the superordinate clause. In English, this dependent status is made syntactically explicit by the subordinator, as in (27) and (28), in which only the conjunctions *although* and *though* indicate that their complement clauses are subordinate to the matrix clause.[21]

(27) **Although** *he collected only 603 votes*, he remains undaunted. (ICE-CAN:W2C-020)

(28) Representatives of many different nations camp in the town, **though** *most are French and Spanish.* (ICE-GB:S2B-027)

The subordinate clause is treated as a constituent of the sentence by Quirk et al. (1985: 987), who argue that it is "downgraded to a subclausal unit". This is essentially why a clausal construction can be substituted with a prepositional one (e.g. *although he failed → despite his failure*).

It has been proposed that concessives are characterised by syntactic constraints that set them apart from other adverbials. Since they are of little relevance to the central analyses of this study, those aspects will only be discussed in the following summary. König (2006: 821; cf. König 1994: 679, 1991a: 192, 1988: 149–151) highlights four syntactic properties of concessive constructions, focusing on subordinate clauses:

1. There are no interrogative adverbs for concessives, corresponding to *when* for temporal clauses and *why* for causal clauses;

2. it is not possible to combine concessives with focusing particles like *only, just, especially* (**He answered her only although he didn't want to.*);

3. concessives cannot be focused in a cleft sentence (*It was although X *that* Y.); and

4. it is not possible to focus a concessive by negation or using a polar interrogative.

In illustration of the fourth point, the question *Did he fail the exam because he was unprepared?* can be answered with *No, he failed because the questions were not fair*, while it is not possible to answer the question *Did he fail the exam although he was well prepared?* by saying *No, he failed the exam although he cheated.*

---

[21]German, in contrast, employs a subordinator and verb-final syntax in the subordinate clause.

The question using *although* is not ungrammatical, but – in contrast to the question using *because* – it can only have wide scope and will receive the respective answer.

As argued by König (2006: 821; cf. König 1991a: 191, Crevels 2000: 314), the specific constraints listed above point to a more general constraint whereby concessive clauses "cannot be focused against the background of the rest of the sentence", which is interpreted as a symptom of their lack of syntactic integration into the matrix structure and as an indication that, in certain respects, concessive subordinates behave more like paratactic elements (König 2006: 821).

Concerning the subordinators themselves, Quirk et al. (1985: 998–999) subdivide them into "simple", "complex" and "correlative". Simple subordinators like *although* and *though* consist of a single word, while complex ones consist of several words (e.g. *even though*, which Quirk et al. do not mention in this context, however). Correlative subordinators consist of two markers, one attached to the subclause and one to the matrix clause (cf. Rudolph 1996: 227). Some CCs with a correlative use of markers will be discussed in §3.5 in the next chapter.

2.3.1 Syntactic ordering

This section focuses on the positions of subordinate clauses relative to the corresponding matrix clauses. Subordinate adverbial clauses can occur in initial, medial and final position, and this is the terminology that will be used in this study. Quirk et al. (1985: 1037) also refer to initial position as "left-branching", medial position as "nested", and final position as "right-branching", while Huddleston & Pullum (2002: 779) speak of "front", "central" and "end" position.

According to Huddleston & Pullum (2002: 780), elements in initial position are placed before the subject; elements in medial position are placed before the verb (and after the subject); and elements in final position are placed after the verb. Quirk et al. (1985: 1039–1040) argue that subclauses in final position are easiest to process, while initial and medial clauses are more difficult to process, particularly if the subclause is long or complex (cf. Huddleston & Pullum 2002: 780; see also discussion below). Some authors focus entirely on the difference between initial and final placement (e.g. Chafe 1984: 437, Wiechmann & Kerz 2013: 1, 7, Diessel 2005: 452), which also makes quantitative analyses more straightforward.[22] In this study, the approach followed in the statistical analyses is to treat clause position as a binary variable, with the two categories "final" and "nonfinal" (cf. §6.3.6).

---

[22]Thus, Wiechmann & Kerz (2013: footnote 2) disregard concessive clauses in sentence-medial position, partly perhaps because they fit a binary logistic regression model to their data. Both initial and medial position are coded as "nonfinal".

Altenberg (1986: 21) argues that a preposed (i.e. sentence-initial) subordinate clause has a "grounding" function – that is, it provides background information against which the (more important) information in the main clause is presented. This arrangement of clauses entails more rigorous advance planning on the part of SP/W, and for Altenberg this is the reason why sentence-initial subordinate clauses are more likely in writing, which is characterised by lower time pressure and allows for post-hoc editing (cf. Diessel 2005: 452). By contrast, conversation is characterised by planning in "real time", with "locally managed" units (including main and subordinate clauses): "[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: 21).[23]

Quirk et al. (1985: 1036) apply the principle of *resolution* to account for the ordering of clauses in complex sentences, saying that "the final clause should be the point of maximum emphasis" (cf. "communicative dynamism" in Quirk et al. 1985: 1556–1557). They use this concept as a sentence-level equivalent of *end-focus*, which applies at the level of the clause, i.e. as a mechanism that can account for the variable arrangement of subordinate and matrix structures in terms of information structure and focus (cf. Chafe 1984: 440; see also discussion in Schützler 2018c).

A fine discussion of competing motivations in the placement of an adverbial clause relative to a matrix clause is provided by Diessel (2005), who does not, however, deal specifically with concessives. The three factors Diessel identifies are related to processing, discourse-pragmatics and semantics. In support of the first principle, and largely based on Hawkins's (1990, 1992, 1994, 2000) "performance theory of order and constituency", Diessel (2005: 458–459) argues that adverbial clauses in sentence-final position are preferable, from the perspectives of both production and parsing: (i) Since the matrix clause is constructed first, SP/W does not need to make an early commitment to a complex sentence structure and is thus relieved of advance planning; (ii) since the subordinator 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, (iii) no (or at least much weaker) constraints are placed upon a subordinate clause in final position concerning its length (or weight). On this basis, the initial placement of subordinate clauses appears as the marked solution which needs to be motivated.

---

[23]Altenberg (1986: 20–24) uses the term "contrastive sequencing" to refer to the ordering of clauses in contrastive (including concessive) constructions. In the present study, more neutral terms like *syntactic ordering* or *clause position* are used, since the phenomenon is far from being unique to adversative/concessive constructions.

One such motivating factor competing with processing-based constraints is what Diessel (2005: 459–461) calls "discourse-pragmatic forces". Although the two concepts are not exactly coextensive, I will discuss discourse pragmatics in information-structural terms (cf. Chafe 1976, 1984: 440, Lambrecht 1996, Krifka 2008, Brinton & Brinton 2010: 324–329; also cf. Wiechmann & Kerz 2013: 3, 6). For example, in (29) the sentence-initial subordinate clause headed by *although* is placed at the junction of two somewhat differently angled sections of the discourse, establishing an elegant transition between them and anchoring the following passage on the antecedent. With regard to information structure, it is also the case that certain specifics concerning the "initiative" referred to in the italicised subordinate clause have been established in the foregoing discourse. We would therefore expect a strong tendency for subordinate structures of this kind to precede the respective matrix clause.

(29)   The economic and social cost of the robberies prevented in the first two years of the initiative is estimated to have been between 107 and 130m, which exceeds the average annual cost (24.1m per year) of the initiative. **Although** *the initiative itself has ended*, funding has been made available to the ten Street Crime Initiative forces in 2005/2006. (BE06, miscellaneous prose)

Reference to earlier parts of the discourse is particularly obvious if there is what I call an *anaphoric* element, i.e. a demonstrative PRO-form, e.g. *this*, as shown in (30).[24] Wiechmann & Kerz (2013: 6) call this a "bridging" context, because the concessive construction (consisting of main and subordinate clause) is explicitly tied into the earlier discourse.

(30)   A notable feature of this study was the number of patients who died before their third dialysis session, often during or immediately after the first dialysis. **Although** *this group is biased in favour of patients with the most severe disease* it may indicate the stress of acute haemodialysis on a compromised cardiovascular system has an adverse effect. (ARCHER, medical writing, 1985)

Finally, Diessel (2005: 461–465; cf. Diessel 2008) discusses semantic factors that influence clause placement. He argues, for example, that prototypical *if*-clauses are predominantly placed in sentence-initial position: They establish a specific semantic frame for the interpretation of following clauses, namely "if A

---

[24]See the discussion of Wiechmann & Kerz (2013) in §5.1.3.

then B (otherwise C)", and it is implied that the early position of the *if*-clause is needed to enable a smooth processing of the clauses following the *if*-condition. Further, in a rearranged sentence of the form "B if A (otherwise C)", B will initially be interpreted as factual, but then needs to be reinterpreted as hypothetical, which "disturbs the information flow" (Diessel 2005: 462) and is thus not ideal. I would argue that iconicity (which Diessel mainly discusses for temporal and causal clauses) can also account for the typical sequence: The *if*-condition in A needs to be met before B can be realised, so that the natural chronology of events (CONDITION → CONSEQUENCE) finds its correlate in the syntactic arrangement of clauses. The argument for iconicity as a motivating factor can also be made with regard to concessives: As shown in §2.2, many concessive constructions are based upon an underlying *if*-*then* relation. Even though the expected outcome or consequence is suspended (or unrealised), one can hypothesise that the natural sequence (IF → THEN) will be iconically represented by the respective arrangement of clauses. This is perhaps why, as pointed out by Rudolph (1996: 232), examples in theoretical discussions often (and in disagreement with actual usage, as shown in Chapter 9) seem to suggest that the subordinate clause typically precedes the matrix clause.

2.3.2 Clause types

In the present study, two complement types of *although*, *though* and *even though* are accepted as distinct syntactic categories and will be explained and exemplified in detail below: (i) finite clauses, including subjunctives and *though*-inversion, and (ii) nonfinite clauses, including present and past participle clauses, as well as verbless clauses.[25]

Finite indicative clauses in combination with concessive conjunctions – as shown in (27) and (28) above – are the most frequent type in the present study. Subjunctives are extremely rare, with only a single example found in the nine components of ICE.[26] This is shown in (31), while (32) is another example from the BLOB corpus (Leech & Smith 2005):

(31)   Mr Dodds says he is quite sorry, and even shook him by the hand when he said goodbye, which is going a bit far to my way of thinking, **though** *he be a fine upstanding young fellow*. (ICE-GB:W2F-005)

---

[25]Grammatical descriptions of subordinate clauses usually distinguish between finite, nonfinite and verbless clauses (e.g. Quirk et al. 1985: 992, Diessel 2005: 451; cf. Givón 1990: 839). The latter two are treated as a single category in this study.

[26]More subjunctives are found in the extended Brown family of corpora, particularly in Brown, BBrown and BLOB, i.e. in data that are older and/or from AmE (cf. Crawford 2009, Kjellmer 2009, Schlüter 2009).

(32)   It is sometimes necessary to remove the second molar, **even though** *it be sound*, in order to give the wisdom tooth enough elbow room to come through. (BLOB; popular lore)

In both examples, the subjunctive mood presents the content of the subordinate clause as less factual: In (31), the positive personal evaluation appears as somewhat less committed, while in (32) a hypothetical situation is discussed. With verbs in the subjunctive, the meaning of concessives shades more strongly into that of conditionals; as discussed in §2.1.1, the latter are an adverbial category to which concessives are related, or out of which they have developed via secondary grammaticalisation.

In the category of finite clauses, there is a word order phenomenon restricted to the conjunction *though*. In (33a), the AdjP complement *difficult* in the original corpus finding is not in its default post-verbal position – shown in the alternative (constructed) subordinate clause of (33b) – but in a slot not only before the subject, but before the subordinator.

(33)   a.   I do not see the whole system of local government finance grinding to a complete halt during the next two years, *difficult* **though** *those years will be*. (ICE-GB:S1B-034; comma added)

       b.   …**though** *those years will be difficult*.

Within a Generative Grammar framework, Culicover (1976: 166–167) and Radford (1981: 213) call this "*though*-attraction" and "*though*-movement", respectively (cf. also Aarts 1988: 44–45), Biber et al. (1999: 908) refer to the phenomenon as "fronting in dependent clauses", while Huddleston & Pullum (2002: 634) call it "preposing in PP structure" (on the basis of their use of the term *preposition*; 2002: 599–600). In the present study, the phenomenon will be referred to as *though*-inversion. Biber et al. (1999: 909) argue that the main purpose in such inverted constructions is to emphasise the preposed element. While both Culicover (1976) and Radford (1981) refer to AdjPs only, the following three examples show that it is also possible for an NP, AdvP, or an entire (nonfinite) VP to precede the conjunction *though* in a similar way.[27]

---

[27]The NP in the subordinate clause in (34) would require a determiner, if, for example, the clause was re-constructed into the unmarked variant (*though he was a brilliant artist*). Intriguingly, the behaviour of such "fronted" NPs resembles that of NPs preceding postpositional *notwithstanding* (e.g. *bad cough notwithstanding* vs *notwithstanding **his** bad cough*), which could be argued to be an equally marked construction (cf. Schützler 2018c).

(34) But Goldie, *brilliant artist* **though** *he was*, isn't in the international league and a lot of overseas collectors aren't all that interested in slightly-known New Zealanders who are a big deal only in their own country. (ICE-NZ:W2E)

(35) The outlaw, *fast* **though** *he was going* […] would have noticed it. (BBrown; adventure and western)

(36) *Shout at Eichmann* **though** *he might*, the Prosecutor could not establish that the defendant was falsifying the way he felt about Jews. (Brown; popular lore)

Nonfinite and verbless clauses are analysable into the same components (or "functional elements") as finite clauses (Quirk et al. 1985: 992). However, the subject is always missing in nonfinite clauses introduced by a subordinator. This can be seen in the following three examples, in which it is impossible to add a subject to the subordinate clause without adding a finite verb as well:

(37) The job, **though** *lacking a certain prestige*, allowed me to write much of the day […]. (AmE06; humour)

(38) Browning's poem 'The Grammarian's Funeral' is a psychoanalysis of a Renaissance scholar who, **although** *born a poet*, devoted his maturity to the examination of abstruse problems in Greek etymology. (BBrown; belles lettres)

(39) **Although** *a rural and predominantly agricultural area*, no part of the Vale is more than 12 miles from major industrial and urban centres. (LOB; learned and scientific)

In (37), subject (*it*) and finite verb (*was*) are not overtly expressed, and the main verb in the subordinate clause appears as a present participle.[28] In (38), subject (*he*) and finite verb (*was*) are also omitted from the subordinate clause, which hinges upon the past participle *born*. Finally, the subordinate clause in (39) neither contains an overt subject (which would have to be *the Vale*) nor a form of the verb BE; it consists only of the subject complement, which is the complex NP in italics.[29] Examples (37) and (38) further illustrate another property of nonfinite

---

[28]Simple-present and perfective uses of the *ing*-participle (as in *seeing him* vs *having seen him*) are not differentiated in the present study.

[29]The distinction between nonfinite subordinate clauses with or without verbs is a rather fine (and, for some purposes, unnecessary) one, as is the distinction between nonfinite components of finite VPs and subject complements. Compare the surface equivalence of *He was tall*/*He was a solicitor*/*He was waiting*.

subordinate clauses, namely that their subject is typically co-referential with the one in the matrix clause (Quirk et al. 1985: 1005; cf. Givón 1990: 836) – a very strong tendency stated as the "normal attachment rule" (cf. Quirk et al. 1985: 1121, Schützler 2018c). Example (39) is an interesting, if not particularly jarring, departure from that rule: The implied subject of the subordinate clause can only be assumed to be *the Vale*, while, strictly speaking, the overt matrix-clause subject is *no part of the Vale*.

The above descriptions are clearly not an exhaustive account of all aspects relevant in the syntactic description of CCs with subordinating conjunctions. An additional point made by Aarts (1988: 41–43) is that subordinate clauses of concession come in different degrees of complexity. He distinguishes three: a "simple" type with no embedded clauses; a "complex" type that contains additional embedded clauses (e.g. *although he stopped when he saw the obstacle*); and a "co-ordinated" type, in which the concessive marker relates to several independent clauses linked by coordinating conjunctions (e.g. *although the food was bad and the staff were unfriendly*).[30] In the present study, subordinate clauses were not coded for degree of complexity, in order not to inflate the quantitative apparatus necessary for analysis, and also because there appeared to be no theoretical reason for doing so. Three syntactic phenomena will be discussed in some more detail in §3.5: (i) the use of correlative conjuncts, (ii) "overlapping" (or "double") concessives, and (iii) the marker *even although*.[31]

2.4 Summary

This chapter set out by providing the historical context for concessives in general and for the particular conjunctions under investigation in this study. Further, different semantic types of concessives – anticausal, epistemic and dialogic – were discussed. Finally, the syntax of present-day CCs involving the three conjunctions was examined, focusing on the position of dependent structures relative to matrix clauses and the structure of complements within subordinate clauses. While the historical background was provided mainly for the general contextualisation of results in this study, the discussion of semantic and syntactic aspects outlines the range of functional and formal variants on which the subsequent quantitative analyses (particularly in Chapters 9–11) will be based.

---

[30]Two thirds of all clauses in Aarts's data were simple; "considerably fewer" were complex; and "only a handful" (Aarts 1988: 43) were coordinated.

[31]These phenomena are based on what was found in the data. Of course, many other marginal construction types are likely to exist, and may be found in other corpora.

As pointed out above, only anticausal and dialogic CCs will be included in the quantitative analyses that are to follow. This is due to the rarity of epistemic and narrow-scope concessives, as well as to the syntactic inflexibility of the latter, which would considerably complicate statistical analyses and generate results so lacking in robustness that they would likely distract from a meaningful interpretation. Similarly, the syntactic options that are considered in the quantitative approach are an idealised abstraction: Apart from the simplified distinction between finite and nonfinite clauses that complement the three conjunctions, clause positions are analysed using a binary scheme, with only a contrast between final and nonfinal positions.

The next chapter will be qualitative in nature, discussing a number of typical (and a few less typical) corpus examples in illustration of the semantic and syntactic structures that were introduced above.

3 Corpus examples

Most of the examples provided in this chapter are taken from the *International Corpus of English* (ICE), which is the corpus exclusively used in the quantitative analyses in Chapters 7–11 (see §6.1). Some of these stem from varieties not otherwise considered in this study (but see Schützler 2018b), namely US-American English and New Zealand English. Additional examples are taken from the *Corpus of Historical American English* (COHA, Davies 2010), and from the eight corpora I collectively refer to as the *extended Brown family of corpora* (or xBrown, for short; see Baker 2009): the BBrown, Brown, Frown and AmE06 corpora of written American English (comprising data from the early 1930s, the early 1960s, the early 1990s and the year 2006, respectively) and the corresponding BLOB, LOB, FLOB and BE06 corpora of written British English.[1] Examples were selected (i) to illustrate the semantico-pragmatic properties of the three semantic types on the basis of more examples than was possible in §2.2; (ii) to show different combinations of semantic types and conjunctions; (iii) to discuss semantically ambiguous cases that defy a clear classification; and (iv) to show interesting syntactic realisations that do not follow the main patterns outlined in §2.3. The first two aspects will be discussed in §3.1–3.3, while (iii) and (iv) will be discussed in §3.4 and §3.5, respectively.

Most examples in §3.1–3.3 follow the majority pattern, i.e. finite clauses complementing subordinating conjunctions. Examples found in the corpora will sometimes be re-constructed, for example by altering the position of a conjunction and thereby changing the status of clauses (subordinate vs matrix).[2] This kind of permutation can help to show the relatedness of semantic types, particularly regarding anticausal and epistemic CCs, and can thus contribute to a better understanding of how they were classified. Where this applies, original

---

[1]BBrown (sometimes called "Lancaster 1931 corpus") was compiled by Marianne Hundt (2004–2013) at the University of Zurich, and AmE06 was compiled by Paul Baker (2010–2011) at Lancaster University; standard references for the other six corpora are Francis & Kučera (1979; Brown), Hundt et al. (1999; Frown), Leech & Smith (2005; BLOB), Johansson et al. (1978; LOB), Hundt et al. (1999; FLOB) and Baker (2009: 312–316; BE06).

[2]The term *re-constructed* is spelled with a hyphen to highlight that it is used in the sense of 'constructed again, in an altered way' and does not refer to inferred historical forms.

corpus examples will be indexed as "a", while derived/re-constructed examples will rank lower in the index, i.e. as "b", "c", etc. Furthermore, the corpus source of original examples will be stated in brackets, but no such statement will be provided for derived (or re-constructed) examples. The connective in each example will be given in bold print, while its complement will be italicised – a convention I already followed in Chapter 2.

3.1 Anticausal concessives

The following (interrelated) characteristics are considered central to a definition of anticausal concessives, as introduced in §2.2.1: (i) Propositions are connected by a topos, i.e. a presupposed relation of cause and effect; (ii) the topos is based on real-world causality and thus goes beyond the mere assumption of a likely concomitance of circumstances; (iii) the topos – and thus the relation between propositions – is not reversible (one proposition is assumed to result in the other, but not vice versa); and (iv) the cause may be directly or indirectly connected to the effect. These four aspects will be discussed in connection with the examples presented in the following paragraphs.

Examples (40a), (41) and (42) are typical instances of anticausal concessives constructed on the basis of the two conjunctions *although* and *even though*. In (40a), the topos is that growing older is likely to result in greying hair, or, more generally formulated: AGEING → CHANGED APPEARANCE. The cause-and-effect relationship is perceived, even though the precise causes (e.g. lower concentrations of pigment as a concomitant of ageing) may not be fully known or understood. In this case, age is an indirect cause, but it is quite firmly linked to the effect (greying hair) and thus the intermediate chain of direct causes is redundant and does not need to be stated.

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

    b. **Although** *Patience was already greyer-haired than Miriam*, she was eleven years her junior. (Re-constructed into epistemic concessive)

The re-constructed variant illustrates what happens when we invert anti-causal concessives: Example (40b) is of course a meaningful concessive construction, but it cannot be classified as anticausal since grey hair cannot be viewed as a direct or indirect real-world cause of advanced age (*CHANGED APPEARANCE → AGEING). Instead, grey hair may trigger certain conclusions concerning its possible underlying causes, which is why the re-constructed example is best

read as an epistemic concessive. The juxtaposition of examples like these highlights crucial aspects of the difference between anticausal and epistemic concessives, as will be further discussed in §3.2.

In (41), the long continuation of an environmental disaster – in this particular case the 1979 Ixtoc I oil spill in the Gulf of Mexico – will normally lead to more severe damages, which is the general topos underlying this example. The extent to which an event of this kind is harmful depends (among other things) on its duration, but duration certainly does not depend on environmental consequences – the topos is not invertible, and the construction cannot be inverted either without changing its semantico-pragmatic status. Topoi like this one are quite complex in that they involve two conditions: IF the effects of a situation are negative (a kind of prerequisite) AND IF the situation persists for a long time, THEN there will be particularly dire consequences. The processing of sentences like (41) poses few problems, which shows that even complex topoi are accessed quite routinely by language users.

(41) And the Ixtoc blow-out in the Gulf of Mexico – **even though** *it gushed for months* – did less harm than it might have […]. (ICE-GB:W2B-029)

In (42), the fact that someone has left a long time ago will normally be expected to result in the fading and loss of the memories associated with them. More generally, the passage of time has certain, normally expected effects on memory: PASSAGE OF TIME → FORGETTING. Contrary to this topos, SP/W in the example states that they still have some remembrance of a person's face, associated with pleasant sensations, even though that person has left long ago.

(42) **Although** *she has left me for a long time*, the rough sketch of her face still floats on my mind like a beautiful picture. (ICE-HK:W2F-008)

Example (43) is based on a topos whereby achieving one's purpose (in this case completing one's studies) increases the likelihood of departing from a certain location. A (somewhat informal) topos could be MISSION ACCOMPLISHED → DEPARTURE, which is a chain of cause and effect perhaps typical of university students, who are often viewed as highly mobile. The syntactic structure of the subordinate clause in this example is also quite interesting as it does not conform to L1 norms.[3]

(43) **Though** *I have finished my studies* I will stay few [sic] more years here. (ICE-IND:W1B-011)

---

[3]In the data, rare syntactic realisations like this were not assigned to a separate category, which is why some of them are discussed qualitatively in this section.

The following, syntactically rather complex example is best understood in an anticausal reading. The subordinate clause states that the proposal under discussion (a projected PhD program) has been commented upon favourably from various sides and that there are no apparent reasons for a delay in its implementation. The matrix clause states that another program, the Ed.D. ("Doctor of Education"), was planned later and is considerably more expensive, but was nevertheless launched earlier than the PhD.

(44)   Meanwhile, **even though** *our proposal has received both external and internal praise, and neither Chancellor Price nor Provost Sellers has raised substantive questions or justified the delays*, the Ed.D. program – planned after ours and costing far more – is up and running. (ICE-USA, business letters)

The concessive reading is strong and straightforward, although the fact that the PhD is not yet up and running is not overtly stated but merely implied, and although two additional arguments – lesser cost and earlier planning of the PhD compared to the Ed.D. – are provided in the matrix clause, i.e. not where they would conceptually belong. The interpretation of such examples poses no problems, which seems to suggest that meaning-making does indeed happen at the constructional level, i.e. on the basis of all the evidence that is provided, and not via a simple one-to-one comparison of propositions in subordinate and matrix clause. Four topoi could be argued to be effective here, two appearing as coordinated parts of the subordinate clause and another two "outsourced" to the matrix-clause parenthesis: (i) POSITIVE EVALUATION → SWIFT IMPLEMENTATION, (ii) NO OBJECTIONS OR QUESTIONS → SWIFT IMPLEMENTATION, (iii) EARLY PLANNING → SWIFT IMPLEMENTATION, and (iv) LOWER COSTS → SWIFT IMPLEMENTATION. All four could of course be subsumed under a more general topos linking positive attributes (like cost-efficiency and good organisation) to success (i.e. swift implementation), and they can also be argued to constitute an interacting causality chain, with early planning and low costs leading to positive evaluation and fewer questions being asked. The structure in the subordinate clause is an example of Aarts's (1988) category of "coordinated concessives" (see §2.3.2) and finds its correlate in the coordinated structure of the matrix-clause parenthesis (*planned after ours and costing far more*).

Examples (45) and (46) provide further illustrations of topoi operative in anticausal concessives. The first one is straightforward: Nurses are expected to have been exposed to and be aware of all kinds of issues to do with the human body, including sexual ones, which is why the subject's ignorance of condoms comes as a surprise. The topos MEDICAL TRAINING → KNOWLEDGE OF BODILY ISSUES would then include not only nurses, but also doctors, for example.

(45) **[A]lthough** *a nurse*, she didn't know what a condom was. (Frown, press
      reviews)

In (46), contrary to expectation, the removal of a law requiring goods from the
American colonies to be shipped to Ireland indirectly (via English ports) does
not lead to direct trade between America and Ireland. The topos is not entirely
universal but requires some understanding of (Western-hemisphere) trade mech-
anisms and of the relevant historical context, both of which are provided by the
context that is not shown here.

(46) **Though** *this restriction was eliminated in 1731*, Irish trade continued
      throughout the eighteenth century to be primarily with England [...].
      (AmE06, learned and scientific)

The examples of anticausal CCs discussed in this section illustrate some of
the topoi that exist, and highlight on what basis occurrences were classified as
anticausal. However, they can only represent a fraction of possible cause-and-
effect relationships that are stored as part of language users' world knowledge
and can be drawn upon when constructing or decoding concessives.

3.2 Epistemic concessives

Three instances of epistemic CCs and their anticausal re-constructions are shown
in (47–49). This semantic type is much rarer than anticausal and dialogic conces-
sives (see results in Chapter 8). While the notion of the *topos* is typically discussed
in connection with anticausal concessives, it plays an integral role with regard to
epistemic concessives, too. However, there is what could be called an inverted di-
rection of inference: Instead of an expected result or effect, an expected or likely
cause or underlying factor is inferred to motivate the observed outcome.
    In (47a), if someone is optimistic about certain developments, one possible con-
clusion might be that this is due to facts or information of some kind (here: "con-
firmation from Baghdad"). That is, given the observed outcome or "symptom",
one makes inferences concerning the possible underlying causes. The mecha-
nism in the epistemic concessive of the example is based on the dissonance be-
tween inferred cause and actual fact.

(47) a. **[A]lthough** *he was optimistic about the release*, he had received no
          confirmation from Baghdad. (ICE-GB:S2B-006)

b. He was optimistic about the release, **although** *he had received no confirmation from Baghdad.* (Re-constructed into anticausal concessive)

In (47b), the sentence has been re-constructed into an anticausal concessive based on the same intra-constructional mechanisms as the examples in the previous section. One could argue that there is a single topos underlying both variant constructions, namely POSITIVE SIGNALS → OPTIMISM.

Example (48a) is about Goh Chok Tong, the second Prime Minister of Singapore, and how he grew up at the time of Singapore's struggle for independence from the United Kingdom during the 1950s. Observing someone like him following his relatives to pro-independence rallies would naturally lead to the conclusion that he is generally involved in pro-independence politics, a conclusion that turns out to be false in this case. Along very similar lines as in (47) above, reconstruction into the anticausal concessive in (48b) is relatively easy.

(48) a. [H]e was not really caught up in the struggle for independence like his uncle and aunt, **though** *he followed them to rallies.* (ICE-SING:W2B-001)

b. **Though** *he was not really caught up in the struggle for independence like his uncle and aunt*, he followed them to rallies. (Re-constructed into anticausal concessive)

In an epistemic reading, the sentence in (49a) could be rephrased as follows: 'Mount Abu – a hill station in Rajasthan, India – has a lot to offer to tourists, although one might conclude otherwise, seeing that it is not as well-known as other Indian hill stations'. Example (49b) once again demonstrates how closely related prototypical epistemic and anticausal CCs are, and how easily one can be transformed into the other.

(49) a. Mount Abu, **though** *a lesser known hill station of the country*, has much to offer to tourists. (ICE-IND:S2B-002; commas added)

b. Mount Abu, **though** *it has much to offer to tourists*, is a lesser known hill station of the country. (Re-constructed into anticausal concessive)

Examples (50–52) are further typical examples of epistemic concessives. In (50), speaking of flux encourages the conclusion that some kind of flow has been observed, but it cannot *result* in there being flux, and thus an anticausal reading is not possible. The construction shown in (51) is an equally clear-cut case of epistemic concession. If one encounters a rug that is 5′7″ by 7′ in size (ca. 3.6 m^2),

one might draw certain conclusions as to its functions, but that of a prayer rug is unlikely to be among them, as such rugs will normally – or prototypically, in Western perception – be smaller. Thus, in the example the proposition marked by *although* triggers certain conclusions and inferences which do not agree with the facts. As I have pointed out in Schützler (2018a: 203, footnote 4), this example is not meaningful in contexts where large prayer rugs of this type are in fact used, and of course in societies or communities that know nothing about prayer rugs at all. Finally, in (52), upward-staring, open but unseeing eyes are likely to lead to the conclusion that a person is dead, while the man in the example has merely fainted. Once again, conclusions drawn on the basis of observed evidence turn out to be in disagreement with reality, which is why the entire construction is categorised as an epistemic concessive.

(50) **[A]lthough** *we speak of flux*, there is nothing which actually flows. (LOB, learned)

(51) **[A]lthough** *five feet seven inches by seven feet in size*, it is a prayer rug [...]. (COHA, 1904, magazines)

(52) The man had only fainted, **even though** *his eyes stared upward, open and unseeing*. (COHA, 1961, fiction)

The examples of epistemic concessives in this section illustrate the different mechanisms involved in the construction and decoding of this semantic type. They have in common that, based on some observation expressed in the subordinate clause, certain inferences are made. Those inferences concern states of affairs (including mental states and personality traits) or events that can be interpreted as having caused or at least contributed to the "symptoms" stated in the subordinate component. It was demonstrated that epistemic concessives can in many cases be conceptualised as inverted anticausal concessives and can therefore easily be re-constructed into the latter type. Anticausal and epistemic CCs seem closely related: Both are explicable in terms of a single inferential mechanism, but they differ in the direction of the inference (CAUSE → EFFECT vs EFFECT → CAUSE). Any proposition will trigger inferences about expected consequences and expected causes, but anticausal and epistemic concessives explicitly capitalise on this, emphasising what could be called *progressive* (forward) or *regressive* (backward) *inference*.

3.3 Dialogic concessives

The class of dialogic concessives is perhaps the most heterogeneous one of the three categories employed in this study. In contrast to anticausal and epistemic concessives, propositions that are juxtaposed in dialogic concessives are not linked inferentially. That is, inferences triggered by the proposition in the subordinate clause can of course not be switched off entirely, but they do not relate directly to the matrix-clause proposition.[4] This definition of dialogic concessives ex negativo will be made clearer by the corpus examples in this section, which demonstrate some of the concrete mechanisms at work in this functional type. To repeat the essentials of what was explained in §2.2.3, the two propositions in dialogic concessives provide pragmatically different comments on the same situation in the sense that (i) they both suggest different conclusions or courses of action, (ii) one qualifies or corrects the other (e.g. curtailing its credibility or the authority on which it is made), or (iii) one provides an alternative perspective on the situation described by the other.

The matrix-clause proposition in (53) describes some cricket ground as "surrounded by slag heaps". There is no obvious inferential link between this and the proposition in the subordinate clause, and the association between the two seems quite loose. What the subordinate clause ("I've not visited myself") does, however, is comment on the credibility of the matrix-clause proposition: SP/W are explicit about not having been to the cricket ground themselves; by making the second-hand nature of the information transparent, the message is qualified, and a more reserved interpretation is encouraged.

(53)   And **although** *I've not visited myself*, the cricket ground is surrounded by slag heaps […]. (ICE-GB:S2A-044; comma added)

The concessive is dialogic in the sense that, metaphorically speaking, AD/R needs to negotiate a conflict that exists between propositions, resulting in a compromise solution for the overall pragmatic outcome. Hilpert (2013a: 166) describes dialogic concessives (which he calls "speech-act concessives", following Sweetser 1990) as "mixed messages", which agrees quite well with the example above. In dialogic CCs of this type, with the subordinate component undermining the authority of SP/W, an inversion (via the reattachment of the subordinator to the other clause) is often not feasible.

---

[4]At the end of §3.2 I suggested that a proposition invariably triggers inferences in one or the other direction (*progressively* or *regressively*, as I put it), but concessives may or may not capitalise on this tendency in the way that propositions are fused into a single construction.

In (54), the age of a sacral building (here: Glasgow Cathedral) is given as seven hundred years, which is followed by a comment to the effect that religious activity in the same location goes back even further than that. This changes the pragmatics of the entire construction by further strengthening the sense of antiquity and tradition that is created. The addition of such informational nuances makes more complex and multi-faceted interpretations possible.

(54)   The best parts of this building are seven hundred years old, **though** *there has been worship here for a great deal longer.* (ICE-GB:S2A-020)

Example (55a) follows a relatively common semantic pattern, which could be labelled UNITY IN DIVERSITY.[5] The matrix-clause proposition focuses on differences between Confucianism and Christianity concerning "the ultimate", while the initial subordinate clause highlights the fact that the goal of finding or experiencing this ultimate is something both have in common. The two propositions provide two pieces of evidence on whose basis Christianity and Confucianism can be compared. Inferential trajectories between the two propositions hardly play a role in this kind of construction; rather, the focus is on their dialogic relationship, characterised by reciprocal qualification. In (55b), inverting the status of clauses by attaching the subordinator to the original matrix clause shifts the focus of the statement, but it hardly affects the interpretation of the whole.

(55)   a. [A]lthough *both traditions direct human being towards the ultimate,* Confucianism discovers the ultimate immanent in human being whereas Christianity finds meaning in the ultimate only by transcending human being […]. (ICE-HK:W2A-005)

        b. Both traditions direct human being towards the ultimate, **although** *Confucianism discovers the ultimate immanent in human being whereas Christianity finds meaning in the ultimate only by transcending human being.* (Re-constructed)

Similar to (55) above, (56) is based on a relatively common pattern, which might be labelled QUANTITY VS QUALITY (cf. Footnote 5). Discussing a particular film genre, the proposition in the matrix clause states that during a certain

---

[5]Setting up a typology of such meaning patterns frequently found in dialogic concessives would be worth an independent research effort but goes beyond the scope of the present study. It may also turn out to be a bottomless pit for the researcher, due to the unknown and potentially vast number of such patterns, their culture-specificity, as well as their open-class character, i.e. the tendency for new ones to emerge.

period in the past, many films of this type were produced in Hong Kong. The subordinate clause elaborates that the films referred to in the matrix clause were not on a particularly grand scale – apparently compared to prototypical exemplars of the genre, or to a specific, present-day example.[6]

(56)   Thirty years ago Hong Kong made many such films, **even though** *not a* [sic] *such grand scale.* (ICE-HK:S2B-033; comma added)

In the example, it does not seem possible to argue for an anticausal or epistemic inferential trajectory between the two propositions; the construction as a whole simply presents two pragmatic stances, one pointing in a more positive direction (QUANTITY = "many"), the other serving as a hedge (QUALITY = "not on such a grand scale"). Constructions of this type also illustrate the lack of a clear boundary between concessive and adversative meaning.

In (57), the proposition in the initial matrix clause assures AD/R that their article will be published soon, only to undermine the meaning of *soon* in the following subordinate clause and thus to imply that it might in fact still take a while for the article to appear.[7]

(57)   So, now you can rest assured that the article is appearing soon, **though** *one doesn't know how to define 'soon'.* (ICE-IND:W1B-008)

While the dialogic element in (53) above lies in questioning the authority of SP/W (whose evidence was qualified as being second-hand, not based on personal observation), (57) is dialogic in questioning the authority (or precision) of language itself.

As in (57), the qualifying proposition in (58) follows the matrix clause. The construction as a whole is mainly concerned with the chances of success of a proposed piece of legislation (the "local option proposal").

(58)   A House committee which heard his local option proposal is expected to give it a favorable report, **although** *the resolution faces hard sledding later.* (Brown, press reportage)

The matrix clause opens on an optimistic note, stating that a positive evaluation is expected in the initial stage of the procedure, while the subordinate clause

---

[6]The message may also be that the number of films was still relatively low compared to other countries, and we would need to turn to the context to work this out more precisely. In this case the label presented here (QUANTITY VS QUALITY) would not hold.

[7]Also note the interesting use of V-*ing* in the matrix clause of this example from IndE.

dampens expectations by adding that "hard sledding", i.e. a more critical assessment and perhaps resistance, is to be expected at a later stage. It is quite typical for the tension between the two different pragmatic stances in dialogic concessives not to be resolved; in fact, it is perhaps one of the main purposes of this functional type to involve AD/R in the meaning-making process (cf. §2.2.3).

The situation in (59a) concerns the poet (and novelist) Thomas Hardy who revised his poems many times; this process of potentially far-reaching aesthetic consequences is qualified by saying that it did not result in dramatic stylistic changes. Re-constructing the sentence by moving the subordinator and thus changing the status of clauses (subordinate vs matrix) once again hardly changes the overall pragmatics, as is demonstrated by the variant example (59b). It is also interesting to note that the core elements of one proposition (*he revised*) resurface as the subject of the other (*the revisions*). This kind of resumed topic – regularly realised as a PRO-form (typically *this*) – makes explicit that both propositions are in fact concerned with a single situation.

(59) a. And **though** *in his later years he revised his poems many times*, the revisions did not alter the essential nature of the style which he had established before he was thirty [...]. (Brown, learned and scientific)

b. In his later years he revised his poems many times, **though** *the revisions did not alter the essential nature of the style which he had established before he was thirty* [...]. (Re-constructed)

Example (60a) hinges upon the juxtaposition of two states of affairs at different points in time: A situation (or state of mind) is altered, perhaps by changing circumstances. In this case, a political or ideological position initially held is modified by social events. Even if one does not fully understand what the sentence is about (namely the food riots in Milan, Italy, on 6–10 May 1898), it is immediately clear that the two propositions do not hold together via an anticausal or epistemic relation, but simply contrast an earlier stage with a later one. The dialogic element consists in the demonstration of changeability: By showing that it changed at a later time, the proposition in the subordinate clause of the original example is made less absolute. A convenient label for this particular type of dialogic concessive could be SEQUENTIAL QUALIFICATION or, more simply, MUTABILITY. Once again, it matters little for the functioning of the concessive which of the two propositions is encoded in the subordinate clause, as shown by the re-constructed variant in (60b). On the other hand, changing the ordering of clauses – irrespective of the attachment of the conjunction – could make the

decoding of the message more difficult, since the actual temporal sequence of events would no longer correspond to the ordering of propositions.[8]

(60)   a.   **Though** *initially he felt his role was to resist the rising tide of mediocrity unleashed by modern mass society* [...], the wide-spread food riots of 1898 left a deep impression on him. (FLOB, learned and scientific)

   b.   Initially he felt his role was to resist the rising tide of mediocrity unleashed by modern mass society [...], **though** *the wide-spread food riots of 1898 left a deep impression on him.* (Re-constructed)

Examples like the following one are quite frequently found in the learned (scientific) texts of xBrown. Their somewhat more abstract structure can be paraphrased as EFFECT, BUT NOT STATISTICALLY SIGNIFICANT. The presence of an effect (here: 'higher resource use in the control group') does not necessarily say anything about p-values, so the two propositions are not linked at the anticausal or epistemic levels. What the construction does is present an interesting effect, which is then toned down by adding that it is not significant in statistical terms.

(61)   a.   [R]esource use among intervention patients tended to be lower than that among the control group, **although** *none of these differences was statistically significant.* (BE06, learned and scientific)

   b.   **Although** *resource use among intervention patients tended to be lower than that among the control group*, none of these differences was statistically significant. (Re-constructed)

As should be clear from the examples cited in this section, there is a vast number of general principles or patterns that may create coherence between the two propositions in a dialogic CC (e.g. UNITY IN DIVERSITY or SEQUENTIAL QUALIFICATION; see above). The identification, discussion and cataloguing both of dialogic subtypes of meaning (as discussed in this section) and anticausal topoi (as discussed in §3.1) can point to general cognitive mechanisms and ways in which humans structure their world knowledge. There is a basic relationship between principles in the anticausal/epistemic and the dialogic domains, but I would still suggest that we need different terms to label them. While the concept of the *topos* is well-established in connection with conditional and causal relations (as operative in anticausal and epistemic concessives), I propose to refer to typical, generalised configurations of dialogic propositions as *themes*.

---

[8]See comments on iconicity in §2.3.1.

3.4 Semantic ambiguity

When coding the data for semantic types, categorical decisions were made: Unless they were truly opaque and had to be excluded, examples were classified as one of the three functional types, anticausal, epistemic or dialogic. There were of course a number of functionally ambiguous cases, which tended to lean towards one of the semantico-pragmatic categories but could also have been plausibly interpreted as a different type (cf. Mondorf 2004: 121–122). Some such examples are reproduced in this section. In many cases, the ambiguity is between two functional categories (e.g. anticausal or epistemic), but there are also instances that display three-way ambiguity, i.e. a potential wavering between all three functional types. CCs of this kind are functional shape-shifters that pose certain problems for the quantitative analysis: The forced classification as one of the three types results in a loss of information, since certain concessives may be characterised by precisely this intermediate position between different functional types and the consequent openness to different interpretations. On the other hand, the inclusion of different degrees of ambiguity (and thus more categories) in the analysis would give rise to considerable complications for the quantitative analysis and the interpretation of results.

Example (62) can be read in three different ways, depending on whether we regard the proposition that is negated in the subordinate clause (namely complete agreement among members) as (i) a prerequisite of the matrix-clause proposition ("official position of the Society of Friends"), as (ii) evidence pointing to the matrix-clause proposition as an underlying cause or motivation, or as (iii) a modification or qualification of the matrix-clause proposition. In the analysis, the dialogic reading was given precedence in cases like this.

(62)   **Although** *not shared by all of its individual members*, this has been the official position of the Society of Friends from its inception in the seventeenth century down to the present time. (BBrown, belles lettres)

The three-way ambiguity is perhaps best understood if the respective thinking is paraphrased in a slightly more abstract way. An anticausal reading would result from the assumption that an official position must be shared by all members of the group, and that it is official *because* it is generally shared. An epistemic reading is essentially an inversion of the first scenario and relies on the reasoning that if there is a lack in agreement, this may be (partly) due to the fact that there has not been any official position or policy concerning this point – this is a possible, but perhaps less plausible interpretation. Finally, if in a dialogic reading

general agreement is regarded as less compulsory, a possible paraphrase would be that 'this has been the official (and therefore quite widely shared) position, but it is not shared by all'. The problem in ambiguous cases like this may be that, while there seems to be some relation of cause/condition and effect, it is not easy to assign those functions to the respective propositions. I would argue that the underlying topos is not clearly enough defined, possibly variable, and affected by subjective experience to a greater extent than in other cases; thus, in this case, classification as dialogic is the most conservative path for the analyst.

An example most likely classified as anticausal but also interpretable as epistemic is shown in (63). Again, the direction of the causal (or conditional) trajectory is not quite clear: Someone may become untrue to themself and their readers by mixing with the wrong people (e.g. royalty and celebrities); conversely, becoming untrue to oneself and one's readers may be viewed as a change in attitude prior to (and ultimately resulting in) mixing with the wrong people. Both an anticausal and an epistemic reading seem possible, and the difference essentially depends on whether one's personal belief is that a change in mental state will result in a change of behaviour, or vice versa.

(63)   **Though** *she mixed with royalty and celebrities*, she always remained utterly true to herself and to her readers. (FLOB, press editorials)

In the following example, categorised as anticausal, one might expect someone majoring in English to have a good command of the language to start with – a relatively high proficiency in English would therefore be a prerequisite for taking a major in the subject, and the construction as a whole would be read as epistemic. On the other hand, one might think that taking a major in English will have the effect of improving a student's command of the language. In this case, there would be a conditional or causal relation between the two propositions in the example, and the construction as a whole would be interpreted as anticausal.

(64)   […] I don't speak good English either, **even though** *I'm taking a major in English*. (ICE-HK:S1A-077; comma added)

There are also cases that are difficult to classify altogether. The construction shown in (65a) – again, most likely classified as dialogic – could be argued to be purely adversative in meaning: There is no obvious causal or conditional connection between the two propositions and they have only a weak qualifying effect on each other. The two propositions (describing the legibility of frequent and infrequent words in an experiment) are merely in a relationship of contrast. What we could say, however, is that presenting both parts of the construction makes

the message complete, as it would perhaps not be satisfactory to be told about frequent words only. In this sense, there is a weakly dialogic element in the construction. As shown in the re-constructed variant (65b), one could quite easily substitute *but* or *while* for *although*, which would arguably make the sentence somewhat easier to interpret.

(65)   a. In word legibility tasks, frequent words were found to be as legible as single letters, **although** *infrequent words are less legible than either.* (Frown, learned)

     b. In word legibility tasks, frequent words were found to be as legible as single letters, **but**/**while** *infrequent words were less legible than either.* (Re-constructed and modified slightly)

Example (66) can be interpreted as anticausal or dialogic; the latter would once again be considered the most conservative option. In the anticausal reading one could argue that expectations will not be formed in the first place if one is aware that they are based on simplistic views. In the dialogic reading, the proposition in the subordinate clause qualifies what we know about the subject, Koesler: The matrix-clause proposition makes him look somewhat naïve, while the subordinate clause adds a more positive nuance to this kind of personal evaluation.

(66)   Somehow, **though** *he knew it was far too facile*, Koesler expected all Italians – as well as Poles, Irish, and Hispanics – to be Catholic. (Frown, mystery and detective fiction)

This section has exemplified constructions whose internal semantic structure allows for alternative readings. Making a categorical decision, i.e. opting for what is felt to be the most plausible reading in a given context, inevitably results in some loss of information – after all, it is possible that the frequency of ambiguous constructions is meaningful in itself. However, the complexity of the quantitative component of this study would have increased considerably had such ambiguous cases been included as a separate category, or even separate categories. A question that might need to be addressed independently is whether or not ambiguous constructions can be shown to have a special function in discourse. In other words: Is the juxtaposition of propositions that allow for multiple (and potentially competing) interpretations accidental or intentional and motivated from the context? Questions of this kind, however, are very complex and cannot be answered in the present study. They may well elude quantitative approaches and are perhaps better addressed in qualitative (e.g. discourse-analytical) studies.

3.5 Further notes on syntax

As anticipated at the end of §2.3.2, there are three syntactic phenomena that deserve a brief discussion, even if they are not treated as distinct categories in the quantitative analyses: (i) correlative conjuncts, (ii) "overlapping" (or "double") concessives, and (iii) *even although* as a marginal concessive subordinator.

The first point concerns correlative marking that consists of a subordinator proper and an optional correlative conjunct, each placed in one of the two clauses that make up the CC. This phenomenon is shown in the constructed example (67a), in which the concessive relation is doubly marked by *although* and *nevertheless*. Given certain syntactic modifications (i.e. the creation of two main clauses), it is possible to dispense with the subordinating conjunctions, as shown in the variant example (67b).

(67) a. **Although** *he was only seventeen years old*, he was **nevertheless** one of the best chess players of the age.

  b. *He was only seventeen years old.* **Nevertheless**, he was one of the best chess players of the age.

As Quirk et al. (1985: 1001) argue, the additional conjunct in the matrix clause has an emphatic function, making the adverbial relation stronger or clearer. The use of a correlative conjunct in the matrix clause may also be motivated by a heavy (i.e. long or complex) preceding subordinate clause, providing a particularly strong cohesive tie between sentence parts and supporting intra-sentential coherence. The following example shows the subordinator *although* in combination with *yet* as a correlative conjunct. It seems very likely that the selection of the correlative marker is motivated by the weight of the subordinate clause in medial position.

(68) This luxurious cabin, **although** *entirely novel to her in conception, design, and furnishing*, **yet** had about it something familiar and personal. (BLOB, adventure and western)

Particularly in certain L2 varieties, *but* is sometimes encountered in addition to a subordinator, as in the following two examples from IndE and HKE, respectively.

(69) **Though** *he was found criminal in the eyes of the law* **but** he couldn't convince himself that he is a criminal. (ICE-IND:S1B-017)

(70) [**A**]**lthough** *it seems that I use a lot of time on studying* **but** the result is
      not […] as satisfactory as others think. (ICE-HK:S1A-038)

L1-oriented language users will most likely try to read (69) and (70) either as
coordinate clauses (with an additional subordinator attached to the first clause)
or as complex sentences with *but* used as a correlative conjunct in the main
clause. The parsing strategy of an L1-oriented AD/R is given a jolt when the
word *but* is encountered. For speakers of the respective L2 varieties of English,
this may of course be quite different.[9]

The next example illustrates what could be described as two overlapping con-
cessive relations. Two subordinate clauses relate to the same matrix clause, one
preceding and the other following it. The example is from published written
material, so this particular double concessive construction must have been con-
sciously planned.

(71) **Although** *the wing structure was only partly supported*, it is believed that
      the wing as a whole was capable of a flapping motion, **although** *soaring
      and gliding was probably the main mode of flight*. (ICE-NZ:W2B-023)

The first one of the two overlapping CCs is best treated as an anticausal con-
cessive, since weak structural support of a wing would normally not result in the
kind of belief stated in the matrix clause. The second part suggests a dialogic read-
ing: "a flapping motion" is indicated as a possibility, but the opinion is expressed
that this was not what the wing was mainly used for. No causal or conditional
trajectory – and thus no topos – operates between the two propositions nearer
the end of the sentence. The construction as a whole quite efficiently first sets the
scene for the matrix-clause proposition, which is then qualified by another sub-
ordinate clause. Examples like this show that concessive constructions may go
some way beyond the simple juxtaposition of two linked propositions.[10] Shared
world knowledge and topoi hold the CC together and make it interpretable, even
if there is great flexibility regarding its syntactic formation.

An interesting if very rare complex marker of concession is the conjunction
even although. In a footnote, Aarts (1988: 41) discusses a single occurrence he
found not in his data but in a letter written "by a Scottish friend"; accordingly, he
is not sure whether this is a feature of Scottish English or simply an idiosyncrasy.

---

[9]Another phenomenon that highlights the problematic and variable status of seemingly straight-
forward connectives is sentence-final *but* (cf. Mulder & Thompson 2008, Mulder et al. 2009,
Hancil 2014, Izutsu & Izutsu 2014), which is also listed as feature no. 211 in eWAVE (Kortmann
et al. 2020).

[10]See also (44) on p. 38 for an interesting, complex case.

A single instance of *even although* was also found in ICE-GB, reproduced as (72), and it was possible to establish the identity of the speaker as Martin O'Neill, a Scottish Member of Parliament. I also came across *even although* in a novel by Scottish writer Peter May, reproduced as (73).[11]

(72) **Even although** *for a moment or two he perhaps enthused the street crowds with the idea that he stood up to the infidels and stood up to the West and to the Americans*, the fact is that he personally has probably killed more Muslims than any other person this century [...]. (ICE-GB:S1B-036; comma added)

(73) I knelt at his head, and kissed him, and prayed for his soul, **even although** *I was no longer certain that there was a God out there.* (May 2012: 383–384)

Since Aarts speculates about a possible Scottish association of this complex connective, and since the only instances that I have come across are from Scottish works of fiction, there may be reason to suspect that this particular form is indeed a Scotticism worth targeting in future research on Scottish Standard English and Scots (cf. Schützler et al. 2017).

3.6 Summary

This chapter presented a selection of corpus examples of concessive constructions. Apart from offering a resource of authentic usage events for future work, it contributes to a better understanding of the concrete semantic mechanisms at work in CCs, which would remain entirely abstract if only the quantitative aspects of the present study were considered. Concerning relations between propositions that are juxtaposed in CCs, the chapter also pointed to certain recurrent semantic patterns, which I conventionally call *topoi* if there is an identifiable causal or conditional link between propositions (in anticausal or epistemic CCs) and *themes* if propositions find themselves in a less narrowly defined, qualifying or corrective relation (in dialogic CCs). While the former term is well-established in the literature, the latter was newly proposed in this chapter. Although it would certainly be worthwhile to work towards a more comprehensive inventory (or typology) of such inter-propositional relations (*topoi* and *themes*), this is clearly beyond the scope of the present study. Furthermore, complications involved in

---

[11]There is at least one more occurrence of *even although* in the same novel (May 2012: 369); this marker has also been independently spotted by another reader in Peter May's *The Chess Men* from the year 2013 (cf. http://languagehat.com/even-although/; last accessed 3 October 2023; however, this blog also suggests that the complex conjunction is generally more widespread).

the semantic classification of cases were highlighted; some of these may be of value as starting points in the development of future (more fine-grained) classification schemes. Finally, non-prototypical syntactic realisations of CCs were discussed. While most of these seem to be of very low frequency, they can play a role particularly in L2 contexts, and they may inform more exclusively syntax-oriented approaches. In sum, Chapter 3 makes explicit what might tend to be lost in the quantitative analyses: CCs are intriguingly complex and in some cases not at all straightforward to categorise, semantically and syntactically. At the same time, language users routinely and effortlessly interpret them, presumably because they rest solidly on shared world knowledge and pragmatic conventions.

4 Dimensions and mechanisms of variation

This chapter discusses three aspects crucial in the context of the present study. Firstly, Construction Grammar is introduced as a theoretical framework used to account for the formal and functional variation of concessives (§4.1). A hierarchical choice model of constructional variation is proposed, in which higher-order properties of a construction have an impact on lower-order properties. These relationships can be employed to predict formal characteristics of CCs. Along with these intra-constructional factors, two language-external factors are introduced in §4.2 and §4.3, respectively: mode of production (speech vs writing) and different geographical or national varieties of English. Analyses of genre that go beyond the general distinction of speech and writing will not be undertaken in the present study.

4.1 Constructions and constructional variation

This section starts with a definition of constructions and Construction Grammar (CxG) in §4.1.1, followed by a discussion of how the CxG framework relates to the usage-based approach, as advocated by Bybee (2001, 2006, 2010), in §4.1.2. These two sections inform the choice model proposed in §4.1.3, which makes special reference to CCs but can be adapted to other constructions as well. The model is cognitively motivated but has direct consequences for quantitative (statistical) models implemented on its basis.

For a brief history of the emergence of CxG, see Östman & Fried (2005); for alternative views of Construction Grammar(s) that partly diverge from the approach taken in this study, see Croft & Cruse (2004: 165–289); concerning CxG and language acquisition, see Goldberg (2003: 222), Tomasello (2005, 2006), and the chapters in part IV of Hoffmann & Trousdale (2013), to name but a few. For a seminal early introduction to the rationale behind CxG, see Fillmore (1988).

4.1.1 Constructions and Construction Grammar

Construction Grammar (or CxG) as defined by Bergs & Diewald (2008: 1) aims to describe grammatical systems in terms of their inventories of constructions at all linguistic levels. Constructions are defined as form-meaning pairings, for example by Goldberg (2003: 219; cf. Langacker 1987, Croft 2005: 274, Trousdale 2012: 168), and in a CxG framework, descriptions of (or theories about) language always need to consider both formal and functional aspects. Croft (2005: 275), for instance, uses the term "element" to refer to any identifiable formal aspect of a construction and the term "component" to refer to any identifiable meaning aspect of a construction. While the labels seem problematic (because easily interchangeable), these two concepts are very much applicable to the present study. A classic example of a construction given by Goldberg (2003: 220) is the "covariational-conditional construction", an instance of which is shown here:

(74)   The more you shout, the less they will listen.

This construction has a characteristic form: Within each of the two elements separated by the comma, the determiner (*the*) does not take a nominal complement but some kind of "comparative phrase" (Goldberg 2003: 220), the two parts are most likely interpreted as clauses but are characterised by non-canonical word order, and they are simply juxtaposed, i.e. not overtly linked by a connective. On the function side, the covariational-conditional meaning of the construction is only accessible from the construction as a whole, i.e. it cannot be derived from the component parts – formal and functional aspects interact and are stored (and used) as a single unit. In contrast to a Generative Grammar approach, for instance, which would either classify the above example as marginal or try to derive it from some underlying main-clause-cum-conditional-clause structure via transformational rules, the constructionist view is "non-reductionist" in assuming that there is nothing beyond (or underlying) the observed form and the associated meaning (Trousdale 2012: 170; cf. Goldberg 2006: 222).

Constructions like the covariational-conditional construction pose obvious problems for traditional syntactic analyses, and they are therefore strong pieces of evidence for a CxG analysis: Non-canonical and unusual forms do not need to be explained as aberrations from a prototypical pattern but can be directly motivated from the specific functions they serve. However, constructions that *do* conform to canonical patterns are also captured by the CxG framework, although they are less conspicuous (see §4.1.2). For instance, a simple SVO clause structure clearly qualifies as a (very general) construction, as does a straightforward combination of matrix and subordinate clause. Thus, while first insights

into CxG are most easily generated by the inspection of syntactically striking examples, a general CxG framework must necessarily capture *all* linguistic expressions.

An important aspect of constructions highlighted by Goldberg (2003: 221) is that "different surface forms are typically associated with slightly different semantic or discourse functions" – that is, in a CxG framework one would naturally hypothesise that a difference in form between two expressions is likely to correspond to some difference in function (or meaning). An example (Goldberg 2003: 221) is the difference between ditransitive constructions (S V O_i O_d : *I bought him* X.) and prepositional-object constructions (S V O_d O_{prep} : *I bought* X *for him.*) – the formal difference between the two is argued to correspond to some difference in function or meaning. This function-form relationship is a crucial element in the analyses presented in this book.

In theory, the term *construction* always refers to a schema (e.g. the "ditransitive construction"), while *constructs* or *allostructions* are realisations of constructions, i.e. lexically filled expressions in use (Bergs & Diewald 2008: 5, Cappelle 2006, Fried 2008: 52). There are of course fixed constructions with very few (or no) options as to how to fill individual slots in the schema. If both syntactic frame and lexical content are relatively fixed, the construction is "lexicalised" or "idiomatic"; if the syntactic schema can be relatively freely filled with lexical content, the construction is "abstract" or "productive" (Bergs & Diewald 2008: 1–2; cf. Goldberg 2013: 18) – the latter type is often referred to as simply "schematic". In the present study, the notion of the *subconstruction* also plays a role. In my definition of this concept, and without tying it explicitly into existing CxG frameworks, subconstructions can be located at levels of schematicity that are intermediate between highly general constructions and constructs that are syntactically fully specific and lexically filled. For instance, if we treat anticausal CCs with subordinate clauses as our maximally schematic construction, then CCs with sentence-initial subordinates and CCs with sentence-final subordinates would be subconstructions at a lower level. These are still lexically unfilled, but syntactically more specific than the general schema. At the next level of specificity, we would then identify the conjunction that is used to connect matrix clause and subordinate clause. Finally, the grammatical status of the subordinate clause (finite vs nonfinite) can be included at an even finer level of granularity. The exact (hierarchical) arrangement of such layers will be partly open to debate, however – for instance, one could disagree about whether it is the choice of a conjunction or the ordering of component clauses that ranks higher, or whether one should place the two on the same level.

The choice model that informs the quantitative analyses in this study is an attempt to formalise a framework of subconstructions for CCs. In this framework, information at more general ("higher") functional and formal levels can be used to predict realisations at more specific ("lower") levels. On the one hand, CCs as constructions do in principle allow for all combinations of functions and forms as defined in this study; on the other hand, there are probabilistic ties between the different functional and formal facets of CCs. This reasoning provides the main link between CxG proper and the quantitative analyses presented in the later chapters of this book. However, both the exact sequence of ranked causes and effects in the proposed choice model as well as the idea of a hierarchy itself may be challenged. Ultimately, the question will be whether the approach contributes to a cognitively grounded explanatory model or simply establishes useful correlations between functional and formal facets of CCs. The latter case would be of value in itself but of course theoretically less satisfying.

4.1.2 Constructions and the usage-based approach

Combining CxG with the usage-based approach (Langacker 1987, 1988, Bybee 2001, 2006, 2010, 2013, Phillips 2006) can generate theories concerning both the emergence and the cognitive representation of constructions, as well as the paths along which those representations change through language use. Bybee's usage-based model – particularly in its version that is geared more specifically towards CxG (Bybee 2013, 2001: 171–177) – is appealing in its capacity for taking into account the multi-faceted (or multidimensional) nature of constructions. Bybee (2013: 51) argues that "[c]onstructions, with their direct pairing of form to meaning without intermediate structures, are particularly appropriate for usage-based models." Combinations of linguistic structure and meaning become entrenched as constructions – i.e. they are turned into "processing units or chunks" – if they are frequently encountered in use. This, Bybee says, happens even if they lack the unpredictable (or idiosyncratic) formal or functional behaviour sometimes regarded as a defining characteristic of constructions (e.g. Goldberg 2003). Thus, even fully predictable structures qualify as constructions if they occur frequently enough (Bybee 2001: 173, Goldberg 2006: 5, Trousdale 2012: 170).

According to Bybee (2013: 53–54), linguistic experience is stored in mental categories called *exemplars*, which exist at all levels of language and also pertain to non-linguistic parameters. Each language event will therefore trigger and be connected to different exemplars, e.g. one that best represents its phonetic properties, one that represents its concrete semantics, one that represents the context of production, and so forth. Exemplars are grouped in an *exemplar cloud* when

they store information concerning the same parameter. For instance, different meanings will be stored in exemplars that belong in a semantic exemplar cloud, and the context in which each utterance is made is stored in the respective exemplar of a stylistic/contextual exemplar cloud.[1] In other words: Each exemplar cloud corresponds to one of the relevant characteristics (formal, functional and language-external) needed for a full description of a particular construction; the exemplars contained within each of these, on the other hand, correspond to a possible realisation of the respective characteristic. If a certain construction is encountered frequently, the relevant exemplars (within their respective exemplar clouds) will be strengthened, as will the connections (or ties) between them. Figure 4.1 provides a schematic illustration, which represents categories relevant to this study. Information about a CC is stored in four exemplar clouds: Cloud 1 contains exemplars of the different semantic types; cloud 2 contains exemplars of different clause positions; cloud 3 stores exemplars of the different conjunctions; and cloud 4 contains exemplars of different complement realisations. The grey lines in the figure suggest that there are connections between any exemplar in one particular cloud and all exemplars in the other clouds. One such combination is highlighted. We can think about this model as a compartmentalised representation: In category E1 (semantics), all instances of concessives encountered by the language user are stored by sorting them into (in this case three) subcategories, or exemplars (e.g. anticausal, epistemic and dialogic). The same happens within the formal categories E2–E4. Along with the exemplars in each cloud, the language user stores degrees of interconnectedness between them, across exemplar clouds. That is, in processing a CC encountered in use, the links between the four involved exemplars are triggered along with the exemplars themselves. Frequent triggering of this kind leads to a general strengthening of particular combinations of functional and formal characteristics, which will then be easier to produce and process. In other words, these CCs become strengthened as subconstructions, as indicated by the black lines in Figure 4.1.

By measuring the strength of certain connections between exemplars in the network based on frequency of use, typical constructional patterns can be identified. According to Bybee (2013: 54), the establishment of such links in the exemplar-based model is one way of conceptualising the emergence of constructions as cognitive representations, and it is in such processes that CxG and the usage-based approach come together. Similarly, Fried (2008: 50) in her constructionist approach views grammar as consisting of "networks of partially overlapping patterns organized around shared features", which comes quite close to the marriage

---

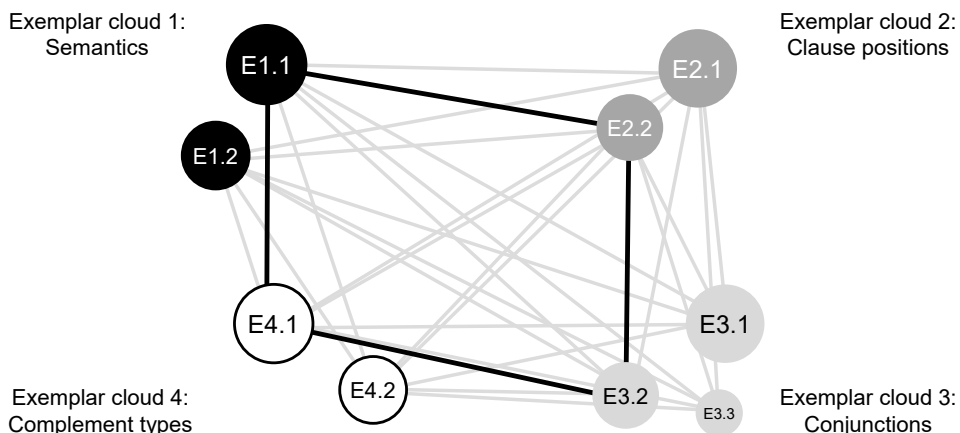[1]Bybee (2013) calls these parameters "criteria".

Figure 4.1: Exemplars and exemplar clouds applied to CCs

of the usage-based approach and CxG in Bybee (2013). Certain elements of the usage-based approach also seem to be implied in publications by Goldberg, when, for example, she refers to a construction as being acquired "on the basis of positive input" (2003: 222) or as a "*learned* pairing of form and function" (Goldberg 2013: 15; my emphasis, OS), or when discussing the concept of "statistical preemption" in the emergence of constructions (2011: 133). The latter appears to be a process rather similar to that involved in the strengthening and weakening of exemplars in the sense of Bybee.

The advantages of this view of constructions as being defined through the strength of ties between exemplars at different levels (semantic, syntactic, etc.) are twofold. For one, it is not merely tolerant of but in fact ideal for the charting of variation, since all exemplars are part of the network, not only the strongest (or most strongly connected) ones. For another, it is quantifiable, since the strength of connections can be measured, either based on the relative frequencies of certain exemplar combinations, or as directional relationships in regression models. The latter approach is taken in this study and will be explained in more detail in the following section. Operationalising connections between exemplar clouds in this (directional, or sequential) way has the disadvantage that the strictly simultaneous view implicit in Figure 4.1 is abandoned: Functional categories have an impact on formal categories, and higher-order formal categories have an impact on lower-order ones, while in a strictly CxG approach, different components would be seen as being on a par with each other. I will argue, however, that the conceptualisation of constructions as tightly integrated sets of ordered choices is useful not only for practical reasons, but also for theoretical ones.

4.1.3 A choice model of constructional variation

In this section, the usage-based model introduced above will be modified by giving the ties between elements in different exemplar clouds a particular direction. While this does not abandon the idea that function and form are inextricably linked (or fused) in constructions, it introduces a certain hierarchical thinking: Higher-order and lower-order characteristics of constructions are assumed to exist, and this ranking can be put to use in theoretical and empirical work. There will be a brief discussion of what I call a *choice model of constructional variation* for English subordinating CCs, building directly upon definitions and descriptions found in §4.1.2 and pointing ahead to the quantitative analyses and their interpretation in the later chapters.

The functional (or meaning) side of a CC is defined by the four semantico-pragmatic types discussed in §2.2, namely (i) anticausal, (ii) epistemic, (iii) dialogic and (iv) narrow-scope dialogic. As has been explained in Chapter 2, only the two most frequent categories (anticausal vs dialogic) will be used in the statistical analysis, but this is irrelevant for the principles outlined here. In this study, then, *function* denotes the relationship between propositions within the construction, or the function of intra-constructional propositions relative to each other. As an alternative to (or expansion of) this relatively local view, which I will call *hermetic*, one could inspect a CC's communicative or discourse function and the relations that hold between it and its wider context of use (see discussion in Chapter 1).

The formal (grammatical) parameters relevant for CCs are threefold: (i) the position of the subordinate clause relative to the matrix clause (cf. §2.3.1); (ii) the connective that introduces the subordinate clause; and (iii) the internal syntactic structure of a subordinate clause (cf. §2.3.2). In the quantitative analysis, the three-way distinction between initial, medial and final position will be reduced to a binary scheme with the categories *nonfinal* (including medial) and *final*. Concerning the third aspect, two complement types are possible in combination with subordinating conjunctions: finite clauses (including subjunctives and *though*-inverted clauses) and nonfinite (i.e. participial or verbless) clauses. This simplified inventory of distinct form-function combinations thus comprises $n = 24$ categories: 2 semantic types × 2 clause positions × 3 conjunctions × 2 complement types. Accordingly, for each of the two functional types that are included (anticausal and dialogic), the number of possible formal realisations is $n = 12$.

The choice model that informs the quantitative analyses in this study is an attempt to formalise a framework of constructional variation for CCs in English. In the model, information at more general (higher) functional and formal levels

can be used to predict realisations at more specific (lower) levels. The following five assumptions are made:

1. Language users produce (and store) constructions.

2. Constructions are correlated functional and formal properties of language.

3. Functional and formal properties of constructions can be ranked:
   a) Form follows function.
   b) Lower-order formal properties follow higher-order ones.

4. The ranking can be employed for
   a) the identification of subconstructions and
   b) the statistical modelling of constructional variation and change.

5. The production of constructions (and subconstructions) is not categorical but probabilistic.

Assumptions 1 and 2 are in broad agreement with existing CxG approaches, have been discussed in different terms in §4.1.1 and will therefore be taken for granted here. The third assumption is based on the (onomasiological) view that the need to express semantic and/or pragmatic meaning is primary, and the linguistic choices that are made to express it are secondary. Further, it is assumed that broader (or more general) formal choices – e.g. located in superordinate structures or heads, in a traditional sense – take precedence over choices corresponding to traditionally lower-ranking structures – e.g. located in subordinate structures or complements/postmodifications. In concrete terms, and with reference to CCs, selecting a general syntactic grid consisting of an arrangement of matrix and subordinate clause (MATRIX→SUB or SUB→MATRIX) is followed by the choice of the marker that introduces the subordinate clause, which in turn is followed by selecting a specific syntactic type of subordinate clause. Figure 4.2 shows the choice model in a schematic form that contains only categories included in the quantitative analyses.

Constructions are still regarded as unitary concepts, with functional and formal parameters inextricably linked. What the model is additionally meant to supply, however, is a framework for the identification of subconstructions at different levels. Starting from a certain functional (or meaning) category, we proceed to different formal layers: There are two subconstructions at the highest and most schematic level, namely CCs with subordinate clauses in final and nonfinal
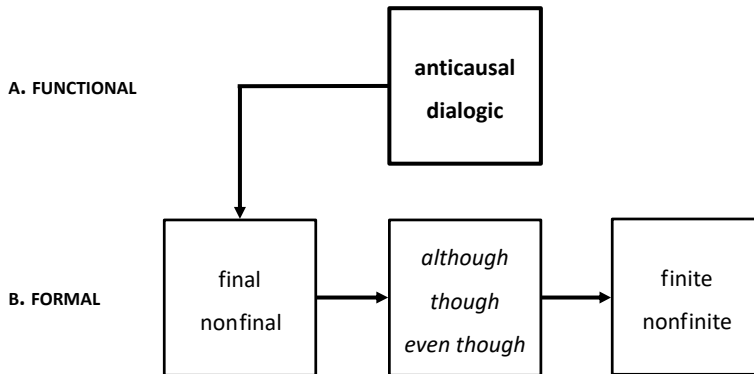
Figure 4.2: A choice model for CCs

position, respectively. Each of these breaks down into three subconstructions at a lower level, distinguished by means of the three conjunctions. At the lowest level, subconstructions are additionally specified for the syntactic class of the subordinate clause.

Figure 4.3 shows the consequences of the choice model for the notion of usage-based CxG: Ties between members of different exemplar clouds are shown only for adjacent levels, once again with one particular combination highlighted in black. The three choices that are made are indicated using arrows, and indexed using the letters A, B and C. We still assume that for any construct, all four parameters – linked by a single path through the four exemplar clouds – must be stored in combination and are not triggered independently. However, the hierarchy of parameters enables us to identify more or less schematic subconstructions, and it will give our quantitative analysis a direction. If this was not the case, it would be hard to decide which parameters to use as predictors, and which as outcomes. These issues will become clearer in Chapters 9–11.

The model is problematic on two counts: (i) It re-introduces traditional grammatical concepts to CxG (e.g. hierarchies, headedness), at least notionally; and (ii) it can at present make no claims regarding cognitive validity. However, it is consonant with the idea that constructions may differ in their degree of schematicity, and it can be used to postulate subconstructions at different levels.
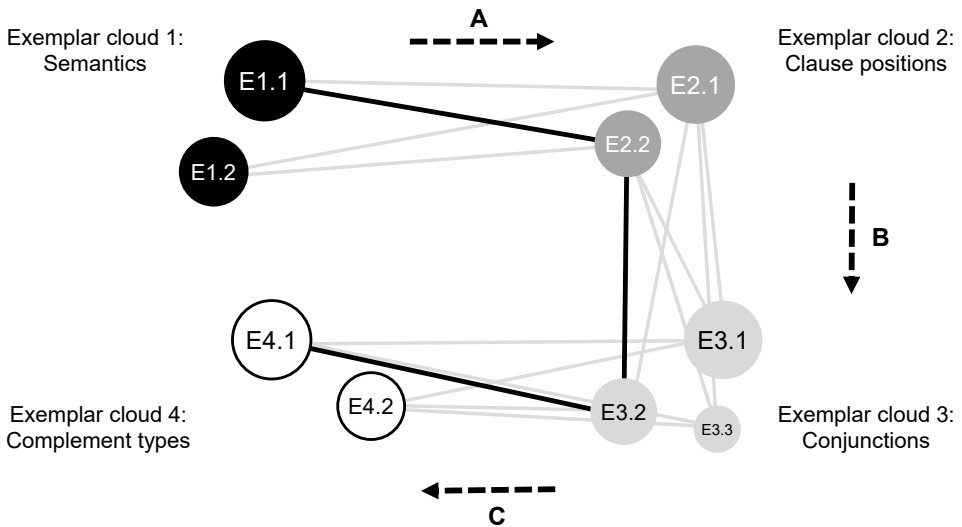
Exemplar cloud 1:
Semantics

A

Exemplar cloud 2:
Clause positions

E1.1

E2.1

E2.2

E1.2

B

E4.1

E3.1

E4.2

E3.2

E3.3

Exemplar cloud 4:
Complement types

Exemplar cloud 3:
Conjunctions

C

Figure 4.3: Merging exemplar clouds and the choice model

4.2 Mode of production

A fine-grained investigation of variation across genres of English is beyond the scope of the present study, which will limit itself to inspections of the two modes of production, speech and writing. While particularly corpora from the ICE family are structured so as to enable the comparison of various written and spoken genres or register (see §6.1 and Appendix A.1), the connectives under investigation are not frequent enough to allow meaningful comparisons at finer levels of granularity. Larger corpora that contain spoken and written material (e.g. COHA, BNC) are of limited use in the World Englishes paradigm since they are restricted to the two main reference dialects of the language, AmE and BrE.

According to Chafe (1994: 42–45), prototypical speech is evanescent, relatively quickly and spontaneously produced, and clearly situated concerning place, time and interlocutors; writing, on the other hand, is produced more slowly than speech, takes a more permanent form and may be edited and revised (cf. Linell 2005: 21). Furthermore, it is desituated, i.e. less clearly tied to a particular temporal, local or circumstantial context. Another basic difference between the two modes of language production is highlighted by Fowler (1991: 59), who associates printed language in particular with "formality and authority" and speech with "informality and solidarity". Fowler also acknowledges that text types written in one medium may assume certain characteristics of the other, so that prototypical writing and prototypical speech constitute the poles of a continuum, rather

than discrete categories. Biber (1988: 45–46) cites face-to-face conversation as an example of typical speech and academic expository prose as an example of typical writing. According to him, academic lectures are an example of speech with characteristics of writing, while personal letters could be described as writing with characteristics of speech. The broad division of the ICE corpus into spoken and written sections merges several such ambivalent genres. As higher-level categories, speech and writing in ICE therefore lack in focus and specificity.[2]

Concerning structural (i.e. linguistic) differences between speech and writing, two central dimensions of variation are proposed by Chafe (1982: 38–49; cf. Chafe 1985, Chafe & Danielewicz 1987, Biber 1988: 21): (i) fragmentation vs integration and (ii) involvement vs detachment. One symptom of the fragmentation of speech is the succession of coordinated (shorter) clauses and the consequently relatively low number of subordinate clauses and the connectives that introduce them (cf. Akinnaso 1982: 104) – a finding that is relevant in the context of the present study. Writing, on the other hand, is more integrated. It contains nominalisations, participles, attributive adjectives, prepositional phrases, and dependent clausal structures (certainly including subordinate adverbial clauses, which are not explicitly mentioned by Chafe, however).

Involvement in oral texts can manifest itself in higher text frequencies of first person pronouns, emphatic particles and hedges; detachment in written texts, by contrast, may result in higher frequencies of passives and nominalisations. It can also be hypothesised that involvement correlates with different pragmatic strategies relevant with regard to different semantico-pragmatic types of concessives (cf. §2.2). For instance, according to Chafe (1982: 45–48), involvement may be reflected in "[r]eferences to a speaker's own mental processes" (e.g. thinking, remembering, reasoning, etc.). Such processes are arguably more transparent in epistemic and dialogic concessives, and less transparent in anticausal concessives. Biber's (1988: 47–49) discussion of explicitness (in writing) and implicitness (in speech) points in a similar direction: Writing is explicit in that it overtly encodes assumptions and logical relations in a text; speech, on the other hand, is more implicit, constructing meaning between interlocutors who jointly contribute to the interpretation process – according to Linell (2005: 18), SP/W and AD/R "co-construct interpretations" in conversation. Thus, it could be hypothesised that

---

[2]Koch & Oesterreicher (e.g. 1985) describe many central differences between speech and writing, in fact anticipating much of what was (independently) formulated by Biber (e.g. 1988) and others. As Biber (1988: 24, 36–37) points out, there is no dimension of variation that simply corresponds to the dichotomy "spoken" vs "written"; from a more general perspective, i.e. ignoring finer textual distinctions within each category, there may be as much variation *within* speech and writing, respectively, as there is *between* them.

the incidence particularly of dialogic concessives, with their pragmatically ambivalent character, will be higher in more involved types of text, and thus in spoken registers.

Finally, there is also a crucial difference between speech and writing in language acquisition, which can help to account for corresponding differences in the use of certain constructions in the two modes of production. As Akinnaso (1982: 111) points out, speech is for the most part acquired "naturally", not at school. The same point is made by Linell (2005: 23), who argues that more explicit instruction is involved in learning to write. Acquiring literacy involves what Linell calls "goal-directed study", based on "explicit norms". Such explicit norms are endorsed by language teachers and codified in grammars, usage guides and teaching materials. The presence (or absence) of certain norms concerning the use of concessives in such reference works may thus contribute to explanations of patterns found particularly in L2 varieties, in which English is acquired scholastically to a greater extent.

It is possible to view speech and writing simply as very general high-level genres. Precisely this is done by Miller & Weinert (1998: 17; cf. Chafe 1994: 48). Within the broad genre category of writing, they argue, one can differentiate between various "sub-genres", e.g. literature, business correspondence, company reports and academic books, which may in turn break down into "sub-sub-genres" (e.g. subdivision of literature into novels, plays, poetry, autobiography and diary). This hierarchical view of genre is also reflected in the sampling scheme adopted for the *International Corpus of English*, for instance (cf. §6.1 and Appendix A.1). At the analytic level, however, only the first-order difference between speech and writing plays a role in the present study; nevertheless, genre differences will sometimes be referred to in a more general way. I use the term *genre* in the same sense as Biber & Conrad's (2009) "text variety", i.e. a sort of text that is produced under certain communicative circumstances. In this terminological decision I follow Smitterberg & Kytö (2015: 118; cf. Meurman-Solin 2001: 243, Moessner 2001: 134–135), who use *genre* for "categories of texts that are defined on extralinguistic or text-external grounds". By contrast, Smitterberg & Kytö (2015: 118) use the term "text type" to refer to categories of text that differ on linguistic grounds. Thus, according to them, "the linguistic make-up of the text itself […] does not determine what genre it belongs to". This is very much in line with the predominant approach in studies that make use of ICE components: Different kinds of text (genres) are sampled from different communicative situations to be then analysed in terms of their linguistic structure.

4.3 Varieties of English

Varieties of English are one of the dimensions across which constructions are assumed to vary in the present study. Section 4.3.1 summarises some general and conceptual issues involved in what has been called the "World Englishes paradigm" (Mesthrie 2003), i.e. the investigation of variation and change in the English language against the background of its spread and diversification across the globe. Furthermore, it introduces the varieties that are studied. Relevant models that have been proposed to describe World Englishes and processes involved in their emergence will be discussed in §4.3.2.

4.3.1 General aspects

English is a pluricentric language (cf. Kachru 1988: 3, Clyne 1992, Leitner 1992) spoken in various locations throughout the world, all of which have the potential of developing their own linguistic norms and standards. Ferguson (1982: vii) considers the spread of English across the globe to be "one of the most significant linguistic phenomena of our time", and for Mesthrie & Bhatt (2008: 12–17) it is a defining characteristic of the Modern English period (cf. McArthur 1998: 87). These views are also reflected in the amount of research on World Englishes that has been and continues to be produced.

Three models of English will be discussed in this section: Kachru's (1985, 1988) *Concentric Circles of English* model, McArthur's (1987) *Circle of World English*, and Schneider's (2003, 2007) *Dynamic Model of the Evolution of Postcolonial Englishes*.[3] Traditional terms that play a more or less central role in many discussions are *English as a native language* (L1 / ENL), *English as a second language* (L2 / ESL), and *English as a foreign language* (EFL).[4] Although the analyses in Chapters 7–11 do inspect patterns in individual varieties, their main objective is to assess cross-varietal stability and variation, not to discuss socio-stylistic patterns and their implications for the status of individual varieties. Models like Schneider's (2003, 2007; see below) therefore serve as a general background to this study but are not exploited to the full. Their discussion in this section is accordingly kept relatively short.

Data from $n = 9$ varieties of English are discussed: British English (BrE), Irish English (IrE), Canadian English (CanE), Australian English (AusE), Jamaican English (JamE), Nigerian English (NigE), Indian English (IndE), Singapore English

---

[3] Jenkins (2015: 2–56) provides detailed summaries of several other models of English.

[4] The terms *English as a lingua franca* (ELF) and *English as an International Language* (EIL) and the – sometimes overlapping – concepts they stand for play no role in my study (cf. Pennycook 1994, Modiano 1999, Jenkins 2000, 2007, Seidlhofer 2011).

(SingE) and Hong Kong English (HKE). Table 4.1 lists the following parameters for each variety: (i) L1/L2 status, (ii) variety label, (iii) world region, and (iv) the developmental phase according to Schneider's (2003) Dynamic Model. Information concerning the latter is taken from Schneider (2007; also cf. Schneider 2011). As the table shows, there are four L1 varieties and five L2 varieties, covering six of the eight Anglophone world regions (cf. Kortmann & Szmrecsanyi 2011: 275, Kortmann et al. 2020): the British Isles (BrE, IrE), America (CanE), the Caribbean (JamE), Africa (NigE), South and Southeast Asia (IndE, SingE, HKE), and Australia (AusE). The Pacific and the South Atlantic are not represented in the study.

Table 4.1: Varieties of English in this study

L1/L2	Variety	World region	Phase
L1	BrE	British Isles	5 / n.a.
	IrE	British Isles	5
	CanE	America	5
	AusE	Australia	5
L2	JamE	Caribbean	4
	NigE	Africa	3
	IndE	S and SE Asia	3–4
	SingE	S and SE Asia	4
	HKE	S and SE Asia	3

At a higher level, the arrangement of varieties in the table and in the visualisations of results follows the division into L1 and L2; within each of these sets, geographical principles are applied, with L1 varieties ordered according to distance from Britain (BrE, IrE, CanE, AusE) and L2 varieties arranged from West to East (JamE, NigE, IndE, SingE and HKE).
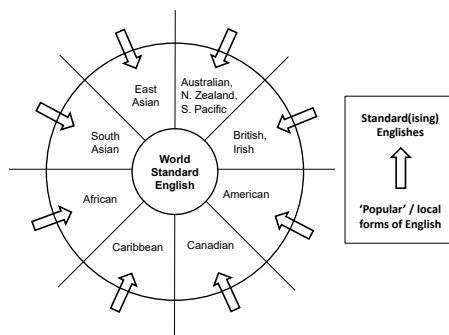
4.3.2 Models of English

The three influential models of English proposed by Kachru (1985, 1988), McArthur (1987) and Schneider (2003, 2007) will be summarised and discussed below. Figure 4.4 gives a first overview, which shows all three models in juxtaposition.

Kachru's is probably the most influential one among models based on circles (Werner 2014: 34, Jenkins 2015: 13). It is motivated by a critique of "a monolingual model for linguistic description and analysis" (Kachru 1985: 11), which prevailed at the time and which is to some extent still reflected in other models (e.g. the

(a) Kachru (1985, 1988)
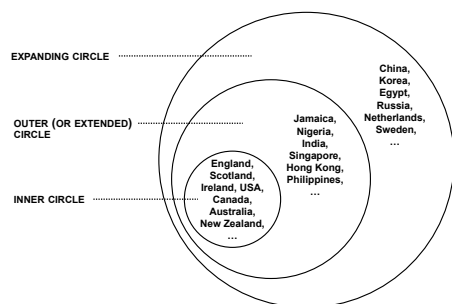
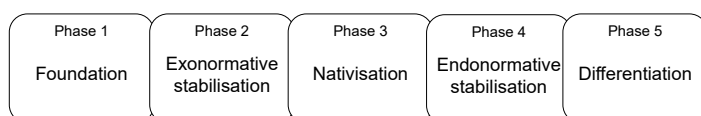(b) McArthur (1987)

(c) Schneider (2003, 2007)

Figure 4.4: Models of English

model by McArthur discussed below; also cf. Görlach 1990). Kachru's *Inner Circle* of Englishes contains varieties which Kachru calls "the traditional bases of English" where it is "the primary language", or L1 (Kachru 1985: 12). The *Outer Circle* comprises varieties of English that have emerged through colonisation – these are what Platt et al. (1984: 3–4) call *New Englishes.*[5] In outer-circle countries, English is a non-native second language (L2), which, however, is given some institutionalised role within the speech community (Kachru 1985: 12–13). Quirk (1985: 4) calls these functions of an L2 within the speech community "internal purposes", as opposed to the "external purposes" of communicating with non-members of the speech community (cf. Greenbaum 1996: 4). The official role of English as an L2 is often, but not necessarily, decreed by political agencies (Platt et al. 1984: 198). Further, L2 English will very often not be the primary language of daily interaction in the home and will therefore first be transmitted through the school system (Platt et al. 1984: 2; cf. Mesthrie & Bhatt 2008: 11). Finally, the *Expanding Circle* comprises those countries or territories which do not have an English colonial background. Here, English is a foreign language not used for internal purposes among members of the speech community (Kachru 1985: 13). The three

---

[5]It is striking how rarely the alternative term "Extended Circle" features in later publications on the subject, although it rather elegantly reduces the terminological distance between the two innermost circles. Perhaps "extended" and "expanding" are too similar and thus too easily mixed up.

circles also differ in the way they adhere to norms or standards (Kachru 1985: 16–17). Inner-circle varieties are *norm-providing* (or *endonormative*), because they are recognised as being used by the native speaker. However, there are considerable differences within this circle in this regard. For instance, Australian and New Zealand English (and probably also Canadian and Irish English) are less widely recognised as norms than BrE and AmE. The Outer Circle is categorised as *norm-developing* by Kachru. Varieties in this circle are both *exo-* and *endonormative*, i.e. outward- and inward-looking for their norms. This implies that one cannot assume a single norm for all levels of the linguistic system, i.e. certain features (or groups of features) may follow an external norm, while others have truly nativised. Even if an outer-circle variety has developed into a norm provider in usage, the new norms will not necessarily be available as a model for language learners, due to a lack of codification (cf. Kachru 1985: 17). In other words, the emerging norms are entirely sociolinguistic and implicit, not pedagogical.[6] Finally, varieties in the Expanding Circle are *norm-dependent*, or *exonormative*; as a general rule, they do not develop norms of their own.

Another circle-shaped model is the one by McArthur (1987: 11), which is shown in Figure 4.4b above. For the sake of simplicity, the figure does not display specific varieties at the periphery of the model (for full details, see McArthur 1987: 11, McArthur 1998: 97). The model is constructed in such a way as to reflect

> the broad three-part spectrum that ranges from the 'innumerable' popular Englishes through the various national and regional standards to the remarkably homogeneous but negotiable 'common core' of World Standard English.

In McArthur's model, regional and national standards and non-standardised varieties form continua in the respective local or national domain.[7] McArthur (1987: 11) is very much aware of some shortcomings of his model, e.g. concerning the relative status of British, Irish, Scottish and English (i.e. Southern British) English, as well as the perhaps overstated difference between American and Canadian English – indeed, despite low speaker numbers, Scottish Standard English

---

[6]This is also true with regard to several L1 varieties. Take, for instance, Scottish Standard English (SSE; cf. Schützler 2015), for which there are few attempts to promote salient and positively evaluated features (for example in pronunciation) and give them a place in education (but see Grant 1914, Abercrombie 1991: 53).

[7]This idea – as well as much of what is said about "World Standard English" by McArthur (1987) – is anticipated in an earlier publication (McArthur 1979: 54–57) much quoted in Scottish English studies, where the concept of a bipolar Scots-English continuum is developed (see also Schützler 2015, Schützler et al. 2017).

may well be a more independent standard than Canadian English, for instance. However, it seems much more interesting to focus on the underlying principles, rather than the exact placement of varieties in the model. For further criticism of McArthur's model, see Mesthrie & Bhatt (2008: 27–28).

McArthur's model also implies that the continua between regional and national standards and their associated non-standard (or "popular") varieties may be extended inwards, resulting in more global continua between the respective world-regional standards and World Standard English. The latter is sometimes associated with certain types of (written) text; for example, McArthur (1987) likens the present-day situation of English to that of classical Latin – whose stability lies in its written form – and refers to "a text-linked World Standard" (McArthur 1987: 10; also cf. McArthur 2003: 56). That is, World Standard English in the sense of McArthur is not codified as such, nor would speakers across the globe have strong intuitions or sentiments about it. Rather, it is "negotiated among a variety of more or less established national standards" (McArthur 1987: 10) whenever the contextual need arises.

The final model to be discussed is Schneider's (2003, 2007: 21–70) *Dynamic Model of the Evolution of Postcolonial Englishes*, usually referred to more simply as the *Dynamic Model*. Schneider's model is dynamic since it assumes five developmental phases through which a postcolonial variety may pass in a certain order, as shown in Figure 4.4c. Varieties can then be characterised according to the stage they have reached. The model assigns a strong role to the speech community and the way its members construct their (postcolonial) identities relative to the (former) coloniser. The model predicts that different kinds of identity construction will result in different kinds of linguistic accommodation (Schneider 2007: 26–29). Like the models discussed above, Schneider's contribution does not claim exclusive validity: The Dynamic Model takes a particular (namely postcolonial) perspective on World Englishes and focuses on aspects not rigorously addressed before. The five phases of the model are briefly summarised in the following paragraph.[8]

In phase 1 (*foundation*), English is brought to a new territory and comes to be used on a regular basis. Settlers and indigenous population have separate identities, the former emphasising the affiliation to the country of origin, the latter regarding themselves as the true and rightful inhabitants of the territory. Cross-cultural communication is relatively limited. There is incipient pidginisation as

---

[8]For a detailed account, see Schneider (2007: 33–55). See also Mesthrie & Bhatt (2008: 32–33), Schneider (2014: 11–12) and Werner (2014) for summaries, as well as contributions in Buschfeld et al. (2014).

well as limited lexical borrowing and a modest degree of bilingualism. In phase 2 (*exonormative stabilisation*), the colonial setting becomes more stable, both politically and linguistically, and colonial administrative structures and the orientation towards the linguistic norms of the colonisers are strengthened. English is spoken more widely. Settlers still view themselves as such, but also perceive a difference between themselves and those at home who do not share the "colonial experience" (Schneider 2007: 37). Many among the indigenous population are beginning to see the benefits of speaking English, and English-speaking indigenous elites emerge. English is beginning to show the first signs of developing into a local variety. In phase 3 (*nativisation*), there are movements towards political and linguistic independence. Cultural, ethnic, economic and linguistic differences between settlers and natives are reduced, and all inhabitants share a sense of belonging to the same territory, despite their different origins. Among the indigenous population, bilingualism is common, and local features begin to stabilise at all linguistic levels. The settler population, on the other hand, divides into two camps: those who readily adopt nativised local features and those who resist. Phase 4 (*endonormative stabilisation*) typically follows the achievement of political independence. Crucially, local identity is now constructed so as to emphasise difference from the mother country. Local forms of English lose their stigma and are widely used to express the new (national) identity; the variety is now endonormative. While local forms are also used in phase 3, in phase 4 such forms are available in more contexts, i.e. not only in the vernacular but also in formal and official contexts. The language variety is perceived as highly homogeneous, even if this need not be supported by the facts of actual usage. Finally, in phase 5 (*differentiation*), the new nation has become politically independent and self-reliant. There is no longer the need to demonstrate linguistic homogeneity, and, accordingly, internal patterns of variation in the speech community emerge more strongly, potentially leading to dialect birth. Identity construction is increasingly driven by social, rather than national factors.

Schneider (2014: 10, 16) is quite explicit about the fact that the Dynamic Model was developed specifically for postcolonial Englishes, i.e. varieties found in Kachru's (1988) Inner and Outer Circles. Testing his own model against English in East Asian expanding-circle contexts, he comes to the conclusion that the model is indeed of limited applicability there (Schneider 2014: 27–28). Based on Schneider's work, Buschfeld & Kautzsch (2017) propose a more flexible and general model they call the model of *Extra- and Intra-territorial Forces* (EIF), which is not discussed here.

4.4 Summary

This chapter has discussed three central topics that provide part of the background for the present study. The theoretical framework of Construction Grammar (CxG) was introduced, with a particular focus on its application to English concessives. The choice model proposed in this context will be instrumental in understanding the approach taken in the quantitative analyses, particularly in Chapters 9–11. Further, two extralinguistic dimensions of variation were introduced: (i) the difference between speech and writing and (ii) the difference between geographical or national varieties of English. In sum, then, the present study treats English concessives as constructions whose different subconstructional levels are assumed to be interconnected and correlated, but which may also be subject to variation induced by different contexts of use. Showing those different aspects in combination will be the task of the analytic chapters.

5 Previous findings and research questions

This chapter sets out by reporting previous research on the concessive constructions that are the focus of the present study (§5.1). The content is arranged in sections that roughly correspond to the sequence of the analytical chapters. Each section proceeds from short summaries of what is reported in the major standard reference grammars of English (Quirk et al. 1985, Biber et al. 1999, Huddleston & Pullum 2002) to discussions of empirical research.[1] Section 5.2 summarises findings in the literature and highlights research gaps (see, in particular, Table 5.1), while §5.3 formulates the concrete research questions and expectations that will be explored quantitatively. These build upon the broader questions formulated in §1.3; their late appearance is due to the fact that they depend on the content of Chapters 2 & 4, as well as the research background summarised in this chapter.

5.1 Previous research

The sequence of content in §5.1.1–5.1.4 roughly parallels the sequence of Chapters 7–11, but direct comparability is in many respects limited: While the main analyses in this book apply negative binomial, binary logistic or multinomial mixed-effects regression models (informed by the constructional choice model introduced in §4.1.3), research in the literature is in most cases based on more basic statistical approaches. That is, text frequencies or percentages/proportions are established, but neither nested data structures nor the interrelatedness of different factors are considered. Moreover, the literature provides hardly any information concerning syntactic types of subordinate clauses, beyond a comment in Huddleston & Pullum (2002). Some of the reported earlier findings will therefore need to be treated with some reservation.

---

[1]Two studies by Burnham (1911) and Quirk (1954) are not included in this chapter because they deal exclusively with concessives in Old English.

5.1.1 Frequencies of conjunctions

The three current major standard grammars of English have rather similar views on the three conjunctions under investigation in this study. Quirk et al. (1985: 1097–1099) regard *although* and *though* as the central markers of concession, the latter being "more informal". *Even though* is treated as an emphatic variant, with the modifier *even* "expressing unexpectedness" (Quirk et al. 1985: 1099). Concerning the basic equivalence of and subtle stylistic difference between *although* and *though*, both Biber et al. (1999: 845) and Huddleston & Pullum (2002: 736) largely agree with Quirk et al. (1985). Biber et al.'s stylistic evaluation of the two conjunctions is based on frequency differences indicating that *though* (including *even though*) is more frequent in conversation and fiction, while *although* is more frequent in academic writing (Biber et al. 1999: 842). They further argue that in preferring *although* to *though* in formal genres, language users may be influenced by the homonymous conjunct *though* (as in *He's quite old, though.*), which is perceived as informal or colloquial (1999: 846; see Schützler 2020b). Interestingly, both Biber et al. (1999) and Huddleston & Pullum (2002) treat *even though* as an intensified variant of *though*, rather than a distinct conjunction in its own right. Biber et al. (1999: 821) further comment that concessive clauses in general – i.e. irrespective of the connective that is used – occur predominantly in written language, which is unsurprising in view of the correlation between writing and the use of complex sentences (see §4.2). On the basis of the information provided by the three major standard grammars of English, one would thus expect *although* and *though* to be the most frequent of the three conjunctions, and one would further expect them to be characterised by different stylistic distributions, even between the two basic modes of production, speech and writing.

Altenberg (1986) presents a study that provides useful descriptive detail on several markers of concession, both in terms of their frequencies in spoken and written BrE and their syntactic positions (concerning the latter, see §5.1.3 below). His work is based on 100,000 words each from the *London-Lund Corpus of Spoken English* (LLC, Svartvik & Quirk 1980, Greenbaum & Svartvik 1990) and the written *Lancaster-Oslo/Bergen Corpus* (LOB; Johansson et al. 1978). Reproducing only results for the three conjunctions relevant in this study, Figure 5.1 shows frequencies in LLC and LOB as well as differences in frequencies between writing and speech, represented as ratios. Raw frequencies are shown in the table on the right of the figure; particularly concerning *even though*, data are sparse and results need to be interpreted with due caution.

In both the spoken and the written material, there is a substantial frequency gap between *though* and *although* on the one hand and *even though* on the other.
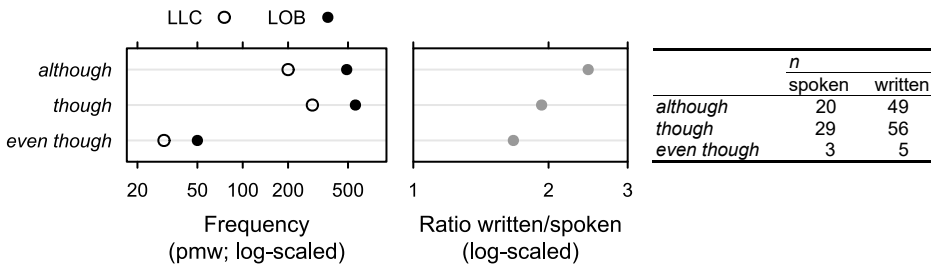
Figure 5.1: Frequencies of concessive conjunctions in spoken and written BrE (Altenberg 1986)

All three conjunctions are more frequent in writing; *although* is most sensitive and *even though* is least sensitive in this regard.[2]

Only slightly later than Altenberg, Aarts (1988) investigates the frequencies of *although*, *though* and *even though* (as well as other concessive markers) based on a sample of 305,000 words from 12 genres in the corpus of the *Survey of English Usage* (SEU; written BrE). As shown in Figure 5.2, *though* is most frequent (*n* = 133; 436 pmw), followed closely by *although* (*n* = 121; 397 pmw); again, *even though* is considerably less frequent (*n* = 16; 52 pmw). The general frequency patterns of the three conjunctions are thus remarkably similar between Aarts (1988) and Altenberg (1986; see above).
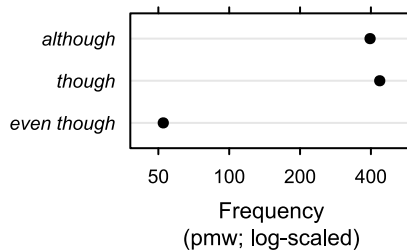
Figure 5.2: Frequencies of concessive conjunctions in written BrE (Aarts 1988)

Regarding stylistic preferences, Aarts (1988: 47–48) finds that *although* is used most frequently in exam essays, medical correspondences, scientific writing, administrative/official language and letters, while it occurs least frequently in journals and non-fiction. By contrast, *though* is stylistically more evenly distributed

---

[2]Note that Altenberg (1986) quantifies the written-spoken difference not as a ratio but using a different index. Further, he does not make any claims concerning *even though* due to its low frequency in his data.

than *although*, i.e. "there are no high peaks in relative frequency for this subordinator", which is interpreted to the effect "that *though* is stylistically less marked than *although*" (Aarts 1988: 50). These findings dovetail nicely with descriptions found in the major grammars. Concerning the third conjunction, *even though*, Aarts (1988: 50) states that its stylistic distribution is even more level.

Schützler (2017) inspects *although*, *though* and *even though* in British, Canadian and New Zealand English (BrE, CanE and NZE), the first two of which are also investigated in this book. Data are taken from the respective components of the *International Corpus of English* (ICE; cf. §6.1). While the study thus foreshadows and is related to the present research, analyses are not regression-based and take a more traditional approach. The main results are shown in Figure 5.3.
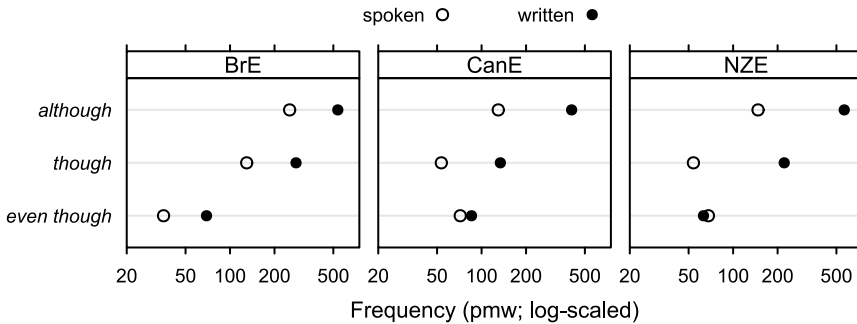
Figure 5.3: Frequencies of concessive conjunctions in three varieties of English (Schützler 2017)

In all three varieties, *although* is most frequent, followed by *though*. There is a clear difference between speech and writing, with higher frequencies in the latter. With regard to *even though*, however, this difference is virtually non-existent in CanE and NZE (Schützler 2017: 177–178). Ratios of writing over speech in BrE (not plotted) are somewhat more regular than in Altenberg's (1986) data (see Figure 5.1), but they are also close to the value of $R = 2$. That is, frequencies in writing are roughly twice those in speech, with $R = 2.1$ for *although*, $R = 2.2$ for *though*, and $R = 2.0$ for *even though*.

Based on the *Corpus of Historical American English* (COHA, Davies 2010), another study by Schützler (2018a) is not so much preliminary but complementary to the present study, using a different corpus and tracing diachronic developments of *although*, *though* and *even though* in AmE from the 1860s to the present day. Figure 5.4 summarises frequencies of occurrence; semantic patterns found in this study are discussed in the next section (§5.1.2).

Frequencies of *although* and *even though* increase over time, while those of *though* decline. Semantic properties and the double function of *though* – i.e. the
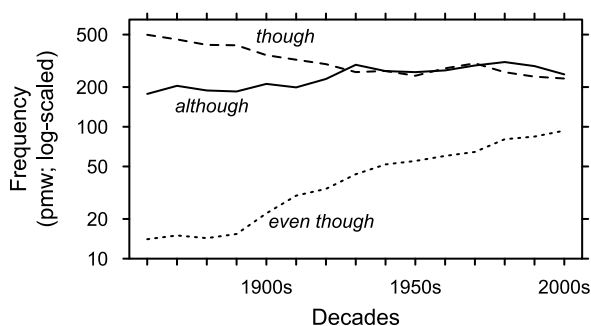
Figure 5.4: Frequencies of concessive conjunctions in COHA (Schützler 2018a: 205)

competition between the use of this form as a conjunction and a conjunct, respectively – are proposed as explanations (see also Schützler 2020b).[3] Based on COHA, the AmE situation at around the turn of the third millennium is broadly comparable to patterns found in the other studies summarised above. Additionally, the diachronic trends can inform hypotheses concerning different patterns in varieties of English, if we assume that the three conjunctions are affected by ongoing processes of grammaticalisation in such varieties.

From a surface perspective, then, the literature suggests that *although* and *though* are used more frequently than *even though*, potentially with *though* taking an intermediate position between the other two conjunctions. Rates of use of *although* and *though* are considerably lower in speech than in writing. Most sources suggest that this is also the case for *even though*, but there is some evidence that this conjunction is used at similar rates in both modes of production, at least in certain varieties. It will be one of the main tasks of the present study to take a closer look at possible underlying functional reasons that may to some extent account for frequency differences.

5.1.2 Semantics

The LLC (cf. Altenberg 1986 above) also forms the basis of Mondorf's (2004) study on gender-conditioned syntactic variation in BrE. One set of outcome variables includes finite adverbial clauses, among them concessives. Mondorf (2004: 85–86) finds that, within the set of adverbials she investigates, only concessives are used more frequently by men than by women. In particular, men use more preposed (i.e. sentence-initial) concessive clauses than women, while there is no

---

[3]The general frequency changes shown in Figure 5.4 progress in a similar fashion in the four broad genres of COHA (Schützler 2018a: 207).

such difference for sentence-final clauses (2004: 99). Furthermore, male speakers use a higher number of concessives that are "propositional" in meaning, which corresponds to what is called the *anticausal* type in the present study (2004: 135–136; cf. §2.2.1). Mondorf concludes that this is because constructions of this type highlight a particularly strong commitment to the truth of a proposition, which correlates with the traditional male domains of authority and power in sociolinguistics (2004: 185–186). As pointed out by Azar (1997), anticausal CCs can strengthen the main clause proposition by anticipating (and defusing) potential counter-arguments, and their use by male speakers is here interpreted as evidence of men's tendency to resort to "linguistic strategies that are least likely to be challenged" (Mondorf 2004: 186). Mondorf's findings contribute sociolinguistically relevant aspects to a discussion of CCs, but in the present study this dimension of variation is not considered.[4]

Hilpert (2013a) studies constructional change in a number of different linguistic structures, including "concessive parentheticals" (155–203). This part of his study is exceptional in presenting a quantitative approach to the intra-constructional semantics of concessives in English (as in Schützler 2017, 2018a). As will be explained below, this semantic aspect is not the main focus of Hilpert's study, which nevertheless provides very important background information for the present investigation. Hilpert concentrates on *although* and *though* (together with other connectives) in written twentieth-century American English, based on data from COHA (Davies 2010). Based on the syntactic behaviour of concessive parentheticals, Hilpert tests two complementary hypotheses concerning their source constructions, namely unreduced concessives (the "reduction hypothesis") and temporal/conditional parentheticals (the "analogy hypothesis"). Formally, Hilpert (2013a: 179) defines concessive parentheticals as reduced clauses that lack a copula and a pronoun as in the following example, in which square brackets indicate the possible reduction:

(75)   **Although** [*it was*] *rare*, family violence did occur. (Hilpert 2013a: 179)

As discussed in §2.3.2, the subjects in both clauses of such constructions (matrix clause and reduced concessive clause) are co-referential, and as a category of comparison Hilpert therefore uses only unreduced (full-clause) constructions in which the subjects of both clauses are also co-referential. His results can therefore not necessarily be assumed to hold for concessives more generally, but

---

[4]Mondorf's (2004) data contain information concerning both the semantics of CCs and the position of clauses, but the relationship (or correlation) between these two parameters is not explored.

rather highlight an interesting correlation of semantics and a particular subject configuration. Moreover, the approach is semasiological, as it quantifies the proportion of semantic types by marker, not vice versa. As shown in Figure 5.5, in Hilpert's data *although* is more often a marker of anticausal concession than *though*, both in full-clause constructions and parentheticals. In both syntactic types, epistemic and dialogic concessives are more frequent in combination with *though*.[5]
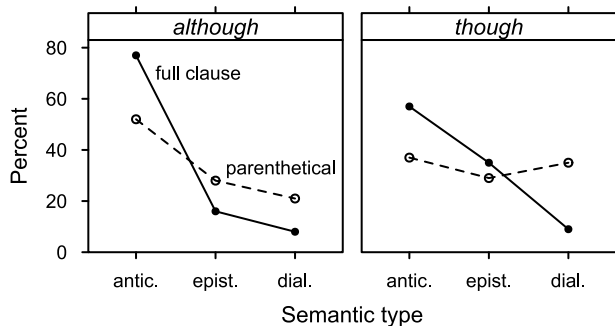

Figure 5.5: Semantic types in full and parenthetical clauses with *although* and *though* (Hilpert 2013a: 189)

The main semantic difference between unreduced concessives and parentheticals appears to be that the latter encode a higher proportion of dialogic concessives while the former associate more with anticausal concessives. Hilpert's study thus suggests that particular semantic types of concessives correlate with particular connectives and with particular syntactic realisations, even though his results are valid only for a subset of constructions, as described above. What is notable in comparison to the results of the present study are the rather high percentages of anticausal and epistemic concessives.

Like Hilpert (2013a), Schützler (2017) takes a semasiological approach to concessive marking; in this case, however, the construction type is not limited to combinations of clauses with co-referential subjects. As shown in Figure 5.6, there is a clear tendency for *although* and *though* to encode dialogic concessives, while *even though* mostly surfaces in constructions of the anticausal type (Schützler 2017: 179–180).[6] The semantic characteristics of concessives employ-

---

[5]In Hilpert's study, the attributes "content", "epistemic" and "speech-act" are used, based on Sweetser (1990). I have taken the liberty of translating them according to the conventions followed in the present study (cf. §2.2).

[6]In contrast to the present study, but in this case in parallel to Hilpert (2013a), Schützler (2017) uses the terms originally introduced by Sweetser, i.e. "content", "epistemic" and "speech-act" (cf. §2.2).

ing the three conjunctions are rather robust, not only across the three varieties but also across speech and writing (not plotted). A very similar pattern is also found in Schützler's (2018a; see §5.1.1) diachronic study of AmE: *although* and *though* predominantly occur in dialogic concessives, while *even though* associates more strongly with anticausal concessives.
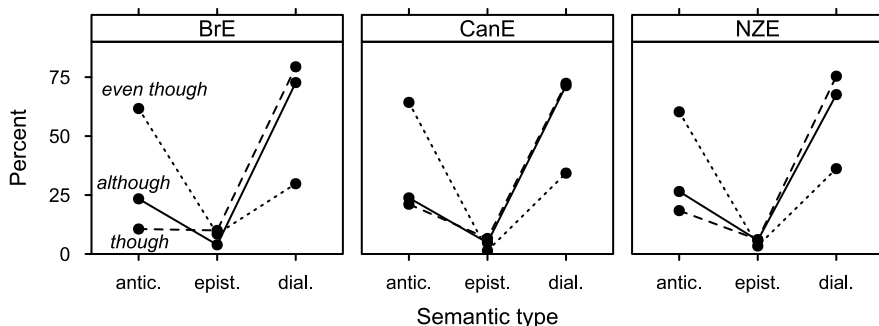

Figure 5.6: The semantics of concessive conjunctions in three varieties of English (Schützler 2017)

Apart from the contribution by Mondorf (2004), only two authors Hilpert 2013a, Schützler 2017, 2018a) have undertaken quantitative analyses that involve the intra-constructional semantics of CCs. The evidence that they present is conflicting: With the exception of parentheticals with *though*, Hilpert (2013a) finds a ranking of semantic types that seems to reflect the historical trajectory of change; that is, the (allegedly original or prototypical) anticausal type is more frequent than the epistemic type, and the pragmaticalised dialogic type is least frequent. In contrast, Schützler (2017, 2018a) finds that the two conjunctions *although* and *though* are most frequently of the dialogic type, followed by the anticausal type, while the opposite is the case for *even though*. Comparability of these two sources is limited because Hilpert focuses on constructions with co-referential subjects in both clauses – which apparently correlate with anticausal meaning – and moreover did not include *even though* in his analysis.

5.1.3 Clause position

Quirk et al. (1985: 1088) claim that – as in conditionals and adversatives – subordinate clauses in CCs tend to be placed before the matrix clause. They give no reason for this pattern, however, and it is quite obviously contra Diessel's (2005; cf. §2.3.1) general theory of clause placement. In contrast to what is claimed by

Quirk et al. (1985: 1088), Biber et al. (1999: 834) find that, across registers, subordinate clauses in concessives are placed *after* the matrix clause in 60% of the cases. Their analysis is limited to finite clauses, however.

Figure 5.7 shows Altenberg's (1986: 22) findings in LLC and LOB (see §5.1.1 above). Subordinate clauses with *though* and *even though* occur more frequently in final position – a tendency that is further strengthened in speech. *Although*, on the other hand, is more likely found in sentence-initial position, but in speech this tendency is less pronounced. Medially placed clauses introduced by the three conjunctions are rare overall. Concerning the general pattern, Altenberg (1986: 22–23) concludes that *although* has a more important "grounding" function for the following discourse than *though* (and *even though*, which he does not discuss due to low numbers). He further argues that different planning strategies in speech and writing are responsible for differences in clause positions between the two modes (Altenberg 1986: 20-22): In speech, there is less advance planning, and a main clause is therefore more often qualified by a postposed subordinate clause (cf. Diessel 2005, as discussed in §2.3.1).
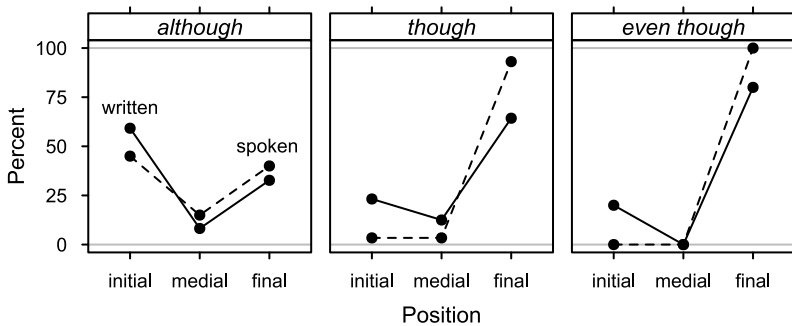

Figure 5.7: Positions of subordinate clauses in spoken and written BrE (Altenberg 1986: 22)

Concerning the position of clauses, Aarts (1988: 43–44) finds that in written BrE *although* is nonfinal in 54% of all cases, while *though* is nonfinal 36% of the time – his findings are shown in Figure 5.8 in greater detail, including the three categories "initial'", "medial" and "final". The ordering of clauses is thus rather similar between the two studies by Altenberg (1986) and Aarts (1988). In both cases, clauses headed by *although* are considerably more likely to precede the matrix clause than clauses headed by *though*. For *even though*, no data on clause ordering are presented by Aarts.

A study by Wiechmann & Kerz (2013) investigates the position of concessive subordinate clauses in the written part of the *British National Corpus* (BNC). Writ-
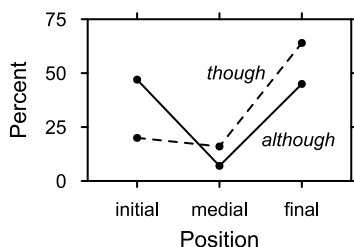
Figure 5.8: Clause positions of concessive conjunctions in written BrE (Aarts 1988)

ten data are used, "as concessive clauses occur predominantly in written registers" (2013: 7; cf. Biber et al. 1999: 821). The study compares $n = 1{,}000$ clauses with the conjunction *although* to $n = 1{,}000$ clauses with *whereas*. In contrast to *although*, *whereas* regularly marks constructions that are purely adversative, rather than concessive, in meaning (cf. Altenberg 1986: 22). For instance, it is difficult to use *whereas* in most of the examples in §2.2 of this monograph. The following summary will therefore focus mostly on *although*.

As independent variables, Wiechmann & Kerz (2013: 3–7) take the following properties of the subordinate clause into account: (i) relative clause length, (ii) finiteness/nonfiniteness, (iii) clause complexity (i.e. constructions with or without an embedded clause within the subordinate clause), (iv) the presence of a "bridging context" that refers explicitly to the preceding discourse with an anaphoric PRO-form, and (v) the choice of the subordinator itself. With regard to *although*, Wiechmann & Kerz (2013: 11–20) find that subordinate clauses are more likely to occur in sentence-initial position if they contain an anaphoric reference to an earlier part of the discourse. Long, complex and finite clauses tend to follow the matrix clause rather than precede it. However, those factors play only "subsidiary roles". Of greater importance is the choice of subordinator: *whereas* tends much more strongly to be placed in sentence-final position. Wiechmann & Kerz (2013) argue that the semantic difference between (concessive) *although* and (adversative) *whereas* motivates differences in syntactic behaviour. The contrast between prototypically concessive and adversative meaning can to some extent be applied to the semantic types of CCs investigated in the present study, too: The dialogic type seems to be closer in meaning to adversativity, lacking the semantic integration (via a topos) characteristic of the other types. While Wiechmann & Kerz (2013) thus identify semantics as an underlying factor, the term (as they use it) refers to rather general categories (e.g. "contrastive/adversative" vs "concessive"), not to the more specific categories established by Sweetser (1990) and used in the present study.

Drawing on written AmE data from COHA (Davies 2010), Schützler (2019) predicts the positions of concessive subordinate clauses (final vs nonfinal) based on several independent variables, among them the subordinating conjunction (*although*, *though* and *even though*) and the semantic type of the construction.[7] Results are shown in Figure 5.9.
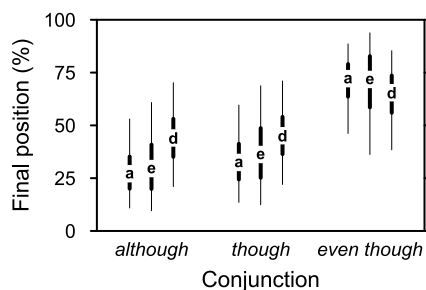

Figure 5.9: Clause positions of concessive conjunctions in COHA (Schützler 2019: 261); a = anticausal, e = epistemic, d = dialogic

The conjunction itself has the greatest impact, with *even though* strongly associating with subordinate clauses in final position. Regarding *although* and *though*, there is a regular effect of intra-constructional semantics: Dialogic CCs tend to be more often realised with subordinate clauses in final position, while preposed subordinate clauses correlate with anticausal semantics.

Finally, Schützler (2020a) also inspects the positions of subordinate clauses, focusing on the conjunction *although* in six L1 and L2 varieties of English.[8] The syntactic behaviour is predicted based on variety status (L1 vs L2), mode of production (written vs spoken) and semantic type (anticausal vs dialogic). Results are summarised in Figure 5.10. There is no systematic difference between L1 and L2 varieties. Both spoken language and anticausal semantics increase the likelihood of subordinate clauses to appear in final position. However, the effect is more moderate in written language. While Schützler (2020a) partly draws on the same data as the present study, the perspective on clause arrangements is rather different: In Schützler (2020a), the choice of marker (*although*) is treated as primary, while in this book the selection of a clausal sequence is regarded as primary in the constructional choice model.

---

[7]Like Hilpert (2013a) and Schützler (2017, 2018a), Schützler (2019) uses different labels for semantic types (cf. Sweetser 1990; see §2.2 above).

[8]The varieties that are inspected are BrE, CanE, NZE, NigE, IndE, as well as Philippine English (PhilE).
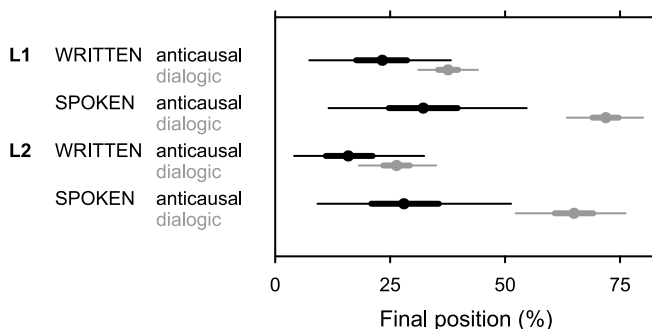
Figure 5.10: Positions of subordinate clauses with *although* in six varieties in ICE (Schützler 2020a)

To sum up the research presented in this section: Speaking for concessives in general, Quirk et al. (1985) and Biber et al. (1999) do not agree on the typical position of subordinate clauses. The empirical studies that were reported indicate that finally placed subordinate clauses are more likely in speech. *Though* and *even though* tend to introduce clauses in final position in some studies, in others it is only *even though* that shows this association. Across the board, however, *although* correlates with clauses in nonfinal position. Further, there is evidence that CCs of the anticausal type are more likely to be found in nonfinal position than dialogic ones.[9] This tendency complicates the assessment of earlier studies like Altenberg (1986) and Aarts (1988), which do not consider the semantic structure of CCs as a potential factor.

5.1.4 Clause types

Information on the relative frequencies of finite and nonfinite concessive subordinate clauses is virtually non-existent in the literature, most likely because there appeared to be no reason to make a special case for concessives in this regard. The only result that can inform the present research is found in Hilpert's study (2013a: 183), where *though* is followed by parenthetical (reduced) structures more often (45% of all cases) than *although* (31%). However, Hilpert's analysis is based exclusively on combinations of matrix and subordinate clauses that share a single, co-referential subject – the prerequisite of clause reduction. That is, the percentage of reduced clauses in the present study will probably be lower, since subjects

---

[9]However, only one of the relevant studies is based on an independent dataset, while the other uses a subset of the data underlying the present study.

in the component clauses of a CC will in many cases be hetero-referential, resulting in irreducible subordinate clauses. Nevertheless, Hilpert's results can be taken to indicate that the shorter connective (*though*) associates with shorter (i.e. reduced) clausal complements.

5.2 Summary and identification of research gaps

The main findings of quantitative studies on the three conjunctions are summarised in Table 5.1, ordered by the three relevant parameters of variation: frequency, semantics and syntax. Asterisks indicate (partial) agreement and mutual support of different studies within the respective category. Superscript daggers indicate disagreement between studies.

Concerning the properties that characterise concessive constructions, clause ordering and frequencies have been investigated in several studies, the latter also with a view to genre-related variation. By contrast, semantic aspects – more precisely: the intra-constructional semantic relations between propositions – have received little attention. Several authors have discussed them in theoretical terms (cf. §2.2), but – with the exceptions of Mondorf (2004) and a few previous investigations by the author of the present volume – the only quantitative study of those aspects is Hilpert (2013a). It is also striking that, except for Schützler (2017, 2020a), concessive markers have not been studied in varieties of English other than BrE and AmE.

Most importantly, the existing research lacks in multifactorial approaches to CCs. As sketched in Table 5.1, disconnected results exist for several dimensions of variation (text frequencies, semantics and syntax), but their interrelatedness (or interaction) remains largely unexplored. Much of the conflicting or inconclusive evidence is therefore likely due to the large number of unknowns in each individual study – i.e. underlying factors that are not operationalised. It is this aspect in particular that the present study will address.

5.3 Research questions and hypotheses

The general research questions for the present study were formulated in §1.3 and will not be restated here. A more precise definition of expectations relies on two components: (i) the constructional choice model introduced in §4.1.3, and (ii) the two extralinguistic factors introduced in §4.2 and §4.3, mode of production and variety status. The diagram in Figure 5.11 shows the expected relationships between external predictors and outcomes at different levels of the construction,

Table 5.1: Concessive conjunctions: Summary of previous research.

Parameter/Source	Findings and comments
FREQUENCY	
*Altenberg (1986)	Frequency ranking: *though > although > even though*; higher frequencies in writing (BrE)
*Aarts (1988)	Frequency ranking: *though > although > even though*; *although*: more frequent in formal styles; *though* and *even though* stylistically less restricted (BrE)
Mondorf (2004)	Higher rates of use of subordinate concessive clauses by men, most strongly if in initial position and/or with proposition-oriented semantics (BrE)
*Schützler (2017)	Frequency ranking: *although > though > even though*; more frequent in writing; *even though* less affected (BrE, CanE, NZE)
Schützler (2018a)	*although/even though*: increasing frequency; *though*: decreasing frequency (AmE)
SEMANTICS	
†Hilpert (2013a)	*although*: anticausal; *though*: mostly anticausal; in parentheticals: weaker preference of anticausal type (AmE)
†,*Schützler (2017)	*although/though*: dialogic; *even though*: anticausal (BrE, CanE, NZE)
*Schützler (2018a)	*although/though*: dialogic; *even though*: anticausal (AmE)
SYNTAX	
†Quirk et al. (1985)	Subordinate clauses precede matrix clauses
†Biber et al. (1999)	Subordinate clauses follow matrix clauses
*Altenberg (1986)	*although*: initial position; *though/even though*: final position; speech: more final position (BrE)
*Aarts (1988)	*although*: nonfinal position; *though*: mostly final position (BrE)
Wiechmann & Kerz (2013)	*although*: long clauses tend to be in final position; anaphoric ("bridging") elements make nonfinal placement more likely (BrE)
Hilpert (2013a)	*though* more likely to take reduced clauses than *although* (AmE)
*Schützler (2019)	*even though*: final position; anticausal: nonfinal position; dialogic: final position (AmE)
*Schützler (2020a)	*although*: higher probability of final placement in speech and in combination with dialogic concessives (BrE, CanE, NZE, NigE, IndE, PhilE)

as well as intra-constructional relationships between parameters. This will be the main point of reference for the analyses implemented in Chapters 7–11 of the book. Note that the model is developed only for anticausal and wide-scope dialogic CCs. That is, epistemic and narrow-scope dialogic CCs are excluded: Both of them are relatively rare, and, in addition, narrow-scope dialogic CCs do not participate in the full range of syntactic variation (see §2.2.4). External factors are represented by triangles (black "W": written language; grey "S": spoken language) and circles (black "L1": first-language/inner-circle varieties; grey "L2": second-language/outer-circle varieties), respectively. Connecting lines between these symbols and components of the construction indicate expected positive correlations (e.g. between spoken language and the selection of *though*, or between L1 varieties and nonfinite clause structures). In addition, the general intra-constructional relations sketched in §4.1.3 (Figure 4.2) are unfolded into more precisely defined expectations, as shown by the connecting arrows. This means that we not only expect higher-order properties to influence lower-order ones, but that we have concrete expectations concerning the way this happens. In the following paragraphs, the reasoning underlying Figure 5.11 will be explained.
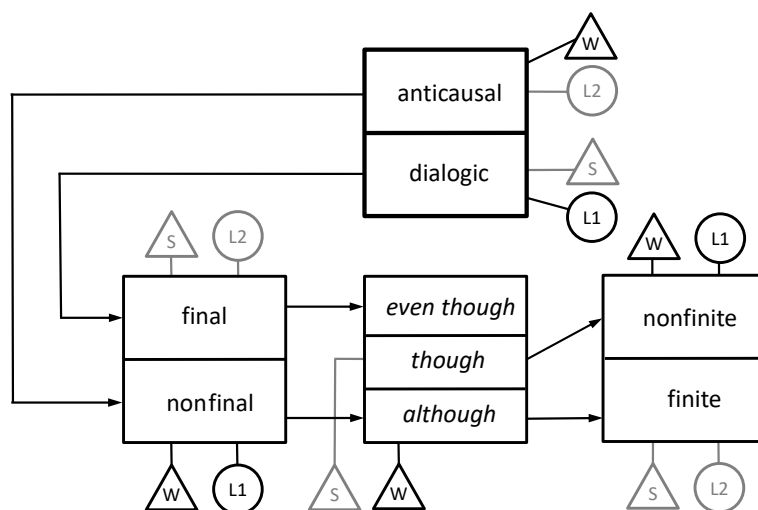

Figure 5.11: A choice model for CCs, including expected correlations

At the intra-constructional level, subordinate clauses in anticausal CCs are expected to be more likely in nonfinal position, since such an arrangement is iconic of the conditional (or cause-and-effect) relation that exists between propositions. In configurations of this kind, the sequence (or dependency) of real-world phenomena finds its correlate in syntactic structure, which is regarded as cognitively

more ideal, both in planning/production and in processing.[10] Dialogic CCs, on the other hand, are not based on such underlying relations and are more often characterised by some sort of (seemingly post-hoc) qualification of the primary statement. They are therefore expected to be more often characterised by subordinate clauses in final position, which has also been argued to be the ideal default configuration, irrespective of adverbial meaning (cf. Diessel 2005; see §2.3.1).

Concerning the relationship between clause positions and the selection of a conjunction, several studies have found that there is a positive correlation between *although* and subordinate clauses in nonfinal position and between *even though* and subordinate clauses in final position. The present study also expects to find these patterns, despite the absence of a firm theoretical basis – except perhaps that a functional specialisation of a marker (or markers) in this regard is generally plausible. Concerning the preferred clause configuration (final vs nonfinal) of *though*, previous research is divided; accordingly, this link is not specified in Figure 5.11. Finally, Hilpert (2013a) provides some evidence of a tendency for *though* to combine with nonfinite clauses more often than *although*, even if his results are only valid for a specific construction type. In this case, it is the conjunction *even though* for which we lack prior information; once again the respective link in the model is left unspecified. Clause positions may also have an effect on the realisation of subordinate clauses as either finite or nonfinite. However, two conflicting hypotheses can be generated. On the one hand, nonfinite (and therefore subjectless) clauses withhold grammatical information, which makes their placement before the grammatically more explicit matrix clause cognitively demanding and therefore less likely. On the other hand, nonfinite clauses will tend to be shorter than finite ones, and according to the principle of "resolution" (Quirk et al. 1985: 1036; cf. §2.3.1) they would be more likely in initial position, leaving the final slot to the heavier matrix clause. No hypothesis is formulated concerning this particular relationship, since the clash of plausible explanations cannot be resolved at this stage.

Mode of production is expected to correlate with the general frequencies as well as the functional and formal parameters of CCs in four ways. Firstly, propositions in dialogic concessives are more loosely connected in that they lack a topos (or underlying causal/conditional presupposition). Often enough, their component parts constitute mutual qualifications, and the construction as a whole comes across as quasi-coordinated in meaning, if not in syntax. Dialogic CCs are therefore considered more compatible with spoken discourse, which puts greater temporal constraints on planning, production and processing. Anticausal

---

[10]There is thus a correspondence between mental presupposition and formal preposition.

CCs, on the other hand, are based on complex inventories of topoi, which need to be accessed by both SP/W and AD/R. In terms of economy and complexity, CCs of this type would therefore be expected to be employed more frequently in writing. Secondly, a positive correlation of spoken language with the final placement of subordinate clauses is expected, while clauses in nonfinal position should associate more with writing. The arguments that underpin these expectations were discussed in §4.2. They are based on the assumption that, from the perspectives of production and processing, subordinate clauses in final position are more straightforward, while subordinate clauses in initial position are cognitively more challenging (see e.g. Hawkins 1994, 2000). Thirdly, based on vague stylistic patterns discussed in the literature, a higher proportion of *although* is expected in writing, while *though* is expected to be relatively more common in speech. The third conjunction, *even though*, is not explicitly discussed concerning its stylistic value in most grammars. Aarts (1988) even makes a point of this marker's equal distribution across different text types, which is why I am reluctant to predict its behaviour across speech and writing. Finally, it is expected that nonfinite clauses should be more frequent in writing than in speech, since they are characterised by a less explicit mapping of surface form onto propositional content and are therefore more easily processed in contexts characterised by lower time pressure. Moreover, using reduced clauses is simply one way of producing more compressed and grammatically less redundant language, which is also more typical of writing (cf. §4.2).

Regarding the correlation between variety status (L1 vs L2) and semantic types, it is expected that anticausal CCs should be relatively more frequent in L2 varieties, while dialogic CCs are more common in L1 varieties. This assumption is based on the general tendency for English to be acquired scholastically in L2 contexts – that is, language users' inventories of constructions and subconstructions are to a larger extent based on tendencies and instructions explicitly codified in grammars and the teaching materials based on them. Even a cursory inspection of grammars like Quirk et al. (1985) reveals that it is predominantly anticausal CCs that are used to exemplify concessive adverbials, and it is expected that this will have at least some effect on language use in varieties that depend on explicit learning to a greater extent. Further, subordinate clauses in final position are expected to be more frequent in L2 varieties: (i) If we accept that finally placed subordinate clauses are the default based on principles of production and processing, then this tendency should be adhered to more strongly in varieties in which contexts of use for English are somewhat more restricted; and (ii) if exposure to English pervades all every-day contexts (as in L1 varieties), there will be more low-level variation and a more flexible handling of syntactic patterns. However,

the factor of scholastic acquisition in L2 varieties may have a contrary effect concerning clause placement. This is because the general tendency for standard grammars (and derived materials) is not only to showcase anticausal CCs, but also to present them with preposed subordinate clauses. There are thus several conflicting hypotheses, and expectations concerning the correlation of L2 varieties with subordinate clauses in final position cannot be formulated with confidence. Next, variety status is not expected to have an effect on the selection of markers. In L2 varieties, *although* and *even though* may be less grammaticalised, and such varieties may therefore be similar to earlier stages of English, as shown in Figure 5.4 for AmE. It seems difficult, however, to position L2 varieties on this kind of historical trajectory, since contexts of acquisition and use are likely to override purely historical factors. Explicitly codified patterns of use may be disproportionally influential, and we would thus expect an even stronger predominance of *although* in such varieties, since this marker is regularly treated as the primary concessive conjunction. These are no more than informed speculations, however, and no clear hypothesis or expectation is formulated regarding the effect of variety status on the choice of conjunction. Finally, I assume that nonfinite clauses will be somewhat less routinely used in L2 varieties: The greater degree of transparency and explicitness that comes with finite clauses may be beneficial in societies characterised by a less pervasive role of English.

The discussion of expected patterns of variation in this section is admittedly multi-faceted and complex. The interpretation of results in Chapters 7–11 will strongly rely on the information presented here, and the reader is invited to refer back to Figure 5.11 when reading the chapters below.

6 Methodologies

This chapter sets out by describing the corpus used for the main analyses in the present study, the *International Corpus of English* (§6.1), followed by an account of procedures adopted in data retrieval, processing and annotation (§6.2). A fairly detailed but general discussion of the statistical approaches central to the analytical chapters is provided in §6.3. The specific formulations of regression models used in Chapters 7–11 are discussed at the beginnings of the respective chapters (in sections entitled "Statistical model").

6.1 The *International Corpus of English* (ICE)

The present study draws on data from the *International Corpus of English* (ICE, Kirk & Nelson 2018), although additional corpus examples presented in Chapter 3 were taken from xBrown as well as from COHA, as explained there, and an example from ARCHER was also cited in Chapter 2. The *International Corpus of English* (ICE) was initiated in 1988 by Sidney Greenbaum at the *Survey of English Usage* (SEU), University College London (Nelson et al. 2002: 2). According to Greenbaum (1988), the project's main objectives were (i) to compile corpora representative of L1 varieties other than AmE and BrE; (ii) to sample L2 varieties of English; and (iii) to build corpora that include spoken and written language (also cf. Greenbaum 1991).[1] However, the potential of ICE for comparative studies of different varieties of English was also a consideration (Greenbaum 1991: 4, 1996: 10). Components of ICE are restricted to L1 and L2 varieties of English; see Greenbaum (1996: 4), for example, who excludes from ICE "English used in countries where it is not a medium for communication between natives of the country", i.e. English in the Expanding Circle (EFL, Kachru 1985, 1988; cf. §4.1).[2]

---

[1]Greenbaum (1991: 4) also points to a possible use of ICE for the monitoring of diversification processes, one function of the corpus being "to preserve the international character of at least written English".

[2]Increasingly, varieties in the Expanding Circle are explored using corpora inspired by ICE (e.g. Edwards 2016, Laitinen 2010). There is also the *International Corpus of Learner English* (ICLE, Granger et al. 2009; cf. Greenbaum 1996 and Granger 1996).

Today, there is an abundance of research that uses the various components of ICE in the World Englishes paradigm (see §4.1).[3]

The design of ICE is documented in several publications (e.g. Nelson et al. 2002: 307–308, Aarts 2011: 347–348). Greenbaum (1992: 171, 1996: 5) refers to the different national versions of ICE as *components*, a convention I will follow. National components of ICE are constructed "along parallel lines" (Greenbaum 1992: 171), and each comprises about 1,000,000 words – 600,000 spoken and 400,000 written (Nelson et al. 2002: 5); this rather limited size of the individual components can be a drawback in the investigation of certain linguistic phenomena. The structure of ICE is often represented as four hierarchical levels: At the highest level, the corpus breaks down into two *sections* (spoken/written), each of which contains two *subsections*. These are dialogues and monologues in the spoken section and printed and non-printed texts in the written section. Subsections are further subdivided into a total of 12 *macro genres*, and at the lowest level there are the 32 specific *genres*. According to the original scheme, speakers and writers sampled for ICE should be at least 18 years old and should have undergone formal schooling in English at least until the end of secondary school, complemented perhaps by a first university degree in L2 countries. A broader selection criterion was to sample "professionals in the widest sense", i.e. "academics, lawyers, politicians, authors, broadcasters, journalists, and business professionals (e.g. managers, accountants)" including students (Greenbaum 1992: 177).

As Greenbaum (1996: 5) points out, there are certain limitations to the representativeness of individual ICE components as well as to their comparability. While the design is the same for all components, the subject matter of texts will naturally not be the same, which may have linguistic effects that are impossible to monitor. Furthermore, speakers and writers are not rigorously balanced according to extralinguistic parameters like sex, age, education or occupation. Finally, certain text categories may be difficult or impossible to obtain in some countries and have to be substituted by related kinds of text.

Greenbaum (1992: 173) states that components should date from the same period, namely 1990–1993. This principle was clearly not upheld, and there is a diachronic dimension both within individual components and between them (cf. Mukherjee et al. 2010: 64, Bernaisch 2015: 64). Many components were compiled

---

[3]See, for example, papers in Hundt & Gut (2012), Seoane & Suárez-Gómez (2016), Hoffmann & Siebers (2009), and in volume 34 of the *ICAME Journal*. See also Aarts's (2011) *Oxford Modern English Grammar*, which uses authentic examples from ICE-GB, very much in the spirit of Quirk et al. (1985: 33), whose reference grammar is partly informed by data from the SEU database and other corpora.

much later (e.g. ICE-Nigeria, Wunder et al. 2010), over a considerably longer pe-riod, or are still under compilation (e.g. ICE-Scotland, Schützler et al. 2017). Early corpora – most notably ICE-GB – might therefore benefit from a second edition that would then be comparable to the newer components.[4]

Markup conventions used for ICE (Nelson 2002a,b) play no role in the present study, since the corpus was searched lexically. Non-corpus material (Nelson 2002a: 8–9, 2002b: 12–13) was excluded from the files before they were searched. This comprises extra-corpus text, i.e. text that simply exceeds the envisaged 2,000 words per text, untranscribed text (e.g. references to tables, formulae or figures in a text), and editorial comments. Non-corpus material was deleted from the files using regular expressions targeting the respective tags in the original files, as documented in some more detail in Appendix A.2.[5] Deleted passages were neither searched, nor were they included in the word counts when measuring the sizes of individual texts (cf. Appendix A.2).

Nine national components of ICE were used in this study, corresponding to the nine varieties detailed in §4.3: Great Britain (ICE-GB), Ireland (ICE-IRE), Canada (ICE-CAN), Australia (ICE-AUS), Jamaica (ICE-JAM), Nigeria (ICE-NIG), India (ICE-IND), Hong Kong (ICE-HK), and Singapore (ICE-SING). They are shown in Figure 6.1.[6] The perspective that is taken is broad, since these nine varieties are scattered across the different world regions or continents. In the study, we can therefore not expect (nor attempt) to reveal patterns that are characteristic of certain kinds of L1 or L2 English, since these are often represented by only one variety. For example, CanE is the only North American English, JamE is the only Caribbean English, NigE is the only African English, and so forth. Further, strong generalisations regarding differences between L1 and L2 varieties are not possible.

Finally, US-American English is not part of the study, since there is no spo-ken section of ICE-USA. The *Santa Barbara Corpus of Spoken American English* (Du Bois et al. 2000–2005) was designed to contain spontaneous conversations of various descriptions, but not to be a true substitute for the complete set of spo-ken genres of ICE-USA, as discussed on the project website.[7] Chafe et al. (1991:

---

[4]At the time of writing, more than 30 years have elapsed since the compilation of ICE-GB, which is roughly the time gap between corpora of the (diachronic) xBrown family (see beginning of Chapter 3).

[5]Thanks are due to Fabian Vetter and Thomas Brunner for their help in preparing the corpus material.

[6]Figure 6.1 is based on a file retrieved from https://commons.wikimedia.org/wiki/File: BlankMap-World_gray.svg; original by user Vardion, transformed into svg-file by Simon Eug-ster; published under a CC BY-SA 3.0 licence: https://creativecommons.org/licenses/by-sa/3.0/.

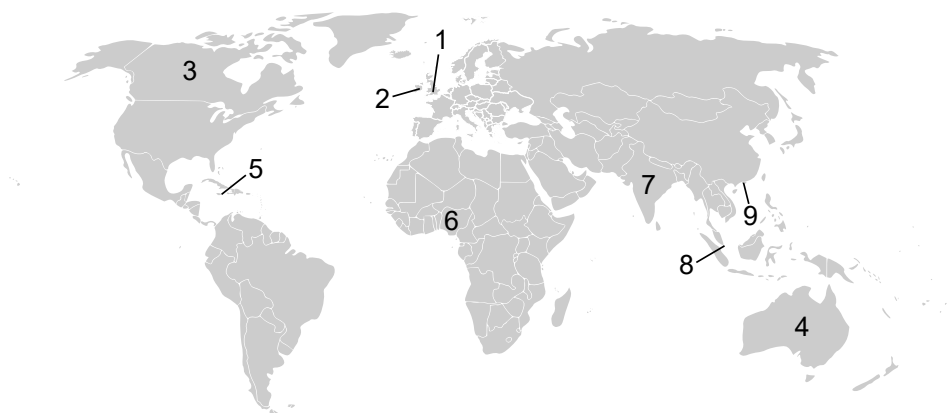[7]http://www.linguistics.ucsb.edu/research/santa-barbara-corpus; accessed 5 October 2023.

Figure 6.1: Varieties investigated in this study. L1: BrE (1), IrE (2), CanE (3), AusE (4); L2: JamE (5), NigE (6), IndE (7), SingE (8) and HKE (9).

64–68) point out that the corpus was originally conceived as an AmE counter-part to the *London-Lund Corpus of Spoken English* (LLC, Svartvik & Quirk 1980, Greenbaum & Svartvik 1990) and was therefore not sampled from a balanced mix of (representative) genres (Chafe et al. 1991: 69). The main text type con-tained in the corpus is face-to-face conversation; other genres comprise phone conversations, on-the-job exchanges, lectures, sermons, story-telling and public meetings or conventions, among others. Direct comparisons with the other ICE components in this study was therefore considered infeasible, and several results in Schützler (2018b) support this assessment.

6.2 Data retrieval and annotation

Components of ICE were used in the form of individual plain text files, with $n = 500$ for the traditional ICE-structure and $n = 902$ for ICE-Nigeria. These were searched using AntConc (Anthony 2018). Non-embedded tags ("<...>") were re-tained in the searches (but not in the word counts; cf. Appendix A.2). The search terms were entered as a list comprising the expressions *although*, *though*, *even though* and *as though*. Adding *as though* to the query made it possible to exclude most instances of this item right from the start. Queries were executed corpus by corpus – that is, nine individual queries and retrievals were run, one for each ICE component. The resulting concordance lists were exported as text files and compiled into a single spreadsheet in Microsoft Excel, which was then taken to the annotation stage. During retrieval, the context window in AntConc had been

set to 300 symbols to the left and right of the search terms; in cases for which this was insufficient for semantic analysis, the full context was reinspected with the "file view" option in AntConc, after a renewed search for the respective occurrence.

Basic coding included labels for variety, text-ID and for the conjunction that was used. Variety labels were added manually when combining the individual output files, text-IDs were part of the output, and the connective (*although*, *though*, *even though*) was found in the KWIC slot ("keyword in context") of each concordance line. Text-IDs were somewhat problematic, since they do not differentiate for variety and may be characterised by diverse formats. That is, filenames like "S1A-027" exist in several corpora, and they may also appear as "s1a-027", i.e. variably with or without capital letters. Since for the mixed-effects regression models (see §6.3.3) it was essential to have one unique text-ID per file, across varieties, I used regular expressions in R to (i) regularise the capitalisation of letters and use of hyphens, and to (ii) add the variety label to each text-ID. This yielded labels such as "GB-S1A-027" which then described unique texts in the set of national components of ICE that was queried.[8]

In the next step, labels for mode of production (spoken/written), genre (e.g. "con" = conversations) and the total number of words in each individual text file were added to each line. The spoken-written distinction was one of the predictor variables, genre was required as a random effect, and the number of words per text was required for the negative binomial model (cf. §6.3.3, §7.2 and §8.1). These pieces of information were added by referring each line in the concordance file to a documentation file that had previously been prepared. This contained the complete information for all individual text files in the components of ICE that were involved: the unique text-ID, the mode of production, the genre and the number of words. Based on the homogenised IDs, there was a match between each of the lines in the concordance file and exactly one line in the reference file, and the relevant information was extracted from the latter and added to the former.[9]

While the parameters described thus far concern the provenance of corpus findings (variety, genre category, text file) and the obvious structural parameter

---

[8]In ICE-NIG, labels took a different form (e.g. "NIG-PHum-001") since this corpus uses explicit genre labels instead of the alphanumerical IDs found in other ICEs (e.g. "W2B") and files are numbered consecutively within each genre category – a practice also followed in the compilation of ICE-Scotland (Schützler et al. 2017).

[9]Fabian Vetter was a tremendous help to me in this regard, as he had documented the number of words per individual text file in several components of ICE for his own work (Vetter 2021, 2022).

of the concessive marker itself, the semantic type of each concessive (cf. §2.2), the position of the subordinate clause (cf. §2.3.1) and the syntactic structure of the subordinate clause (cf. §2.3.2) had to be coded manually. Although this was technically less involved than the process described in the previous paragraph, it was naturally much more time-consuming. During this process of semantic and syntactic annotation, there was some additional disambiguation and exclusion of false positives and items that could not be used for other reasons. For example, (i) instances of *though* were manually inspected and classified as either conjunctions or conjuncts (cf. Schützler 2020b) and were excluded if they were found to be conjuncts; (ii) items were discarded if the context was too fragmented for meaningful analysis, which was considerably more often the case in speech than in writing; finally, (iii) false positives like the ones illustrated in (76–79) were sorted out (all emphases in these examples are my own; OS). In (76), *although* is not used as a grammatical marker but quoted as a linguistic object; in (77) and (78), there are obvious transcription errors (<though> should read <through> and <thought>, respectively); and in (79), there is a false start resulting in a duplicate marker (leading to the exclusion of the first instance of *though*). Other cases not shown here included instances in which it was simply not possible to classify *though* as a subordinating conjunction or a conjunct.

(76) New topics should begin in a new paragraph and there should be a proper link between the two paragraphs and to do this proper words like **Although**, However, etc. (ICE-IND:W1A-013)

(77) Every bric-a-brac hurtled **though** the living room. (ICE-PHI:W2F-005)

(78) They gave little **though** to what he was called in the macrocosm outside. (ICE-HK:W2F-020)

(79) **Though** I **though** *I have worked as both teacher and an administrator* I had more satisfaction as a [sic] administrator than as a teacher […]. (ICE-IND:S1A-026)

After the data had been checked in this way and all valid cases had been annotated semantically and syntactically, the dataset was ready for statistical analysis in R and was exported as a comma-separated file.

6.3 Mixed-effects Bayesian regression models

This section introduces mixed-effects Bayesian regression modelling, the main tool for statistical analysis in this book. After a general introduction to hierarchical data structures and the consequent need for mixed-effects models (§6.3.1),

the rationale behind and the advantages of Bayesian models are discussed (§6.3.2). Further, the three relevant model types are described (§6.3.3), followed by a few words on random effects and centred predictors (§6.3.4), a description of the estimation and visualisation process (§6.3.5) and a definition of all variables used in the models (§6.3.6). All analyses were conducted with R (R Core Team 2021), working within the RStudio environment (RStudio Team 2009–2021).

6.3.1 Hierarchical data structures

Naturally occurring, observational language data almost invariably have a hierarchical structure (see Johnson 2014, Speelman et al. 2018: 2–3, Winter 2020: 232–233, Winter & Grice 2021), which means that data points are not independent but clustered or grouped. In this study, the grouping of observations happens at three levels: (i) the variety, (ii) the text category (or genre), and (iii) the author (or speaker). All three categories constitute language-external, contextual grouping factors: Data points produced in certain contexts or by certain individuals are assumed to belong together and form a group. Although this is not relevant in the present study, language-internal groupings are also common: If, for instance, we are interested in the dative alternation in English, there will be several observations involving the same verb (e.g. *give* or *show*), which introduces a grouping structure on linguistic grounds (cf. Speelman et al. 2018: 2).

Hierarchical structures can be conceptualised as resulting from drawing a "multistage sample" (Hox 2010: 4; cf. Gelman & Hill 2007: 7), where lower-level units are sampled from higher-level units, potentially at several levels (see also the discussions in Section 3.4.1 of Sönning 2020, and Sönning & Schlüter 2022). It is essential that such clustered structures are taken into account when statistically modelling language variation. In this study, the clustering of data by variety is addressed by fitting a separate model for each of the nine varieties. The three grouping levels are thus reduced to two, and, accordingly, there are only two random-effects components in the statistical models, GENRE and TEXT. Each variety, then, contains a sample of genres, each of which in turn contains a sample of texts, in both cases according to the design of ICE. Finally, each text contains a number of individual observations. The structure is fully nested, not crossed (or partly non-nested; cf. Baayen 2008: 260–261), since an individual observation is attributed to one (and *only* one) specific text, and each text belongs to one (and *only* one) genre within the respective variety of English. Figure 6.2 shows the hierarchical data structure in each of the nine varieties in a schematic form.[10]

---

[10]Note that the concrete design of a statistical model is based on active decisions taken by the researcher, which, in turn, are based on their assumptions about the data, as well as their aims concerning generalisability.
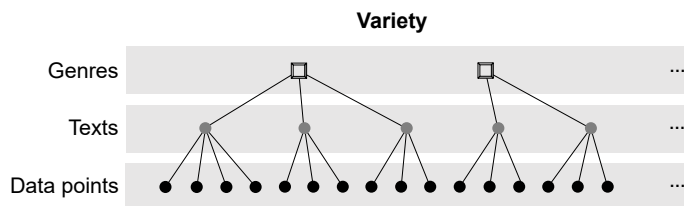
Figure 6.2: Grouped (or hierarchical) data structure in the present study

Two issues may arise if grouped data structures are not taken into account: (i) Point estimates (e.g. percentages of variant constructions or frequencies of markers/semantic types) may be less precise (Sönning & Schlüter 2022) and (ii) our assessment of statistical uncertainty may be overly optimistic or pessimistic. Typically, uncertainty tends to be underestimated for between-cluster effect estimates and overestimated for within-cluster effect estimates (Agresti 2013: 489). Since data points in non-hierarchical models are treated as independent, we effectively inflate our sample size. In the schematic structure of Figure 6.2, let us assume that the first ten data points stem from spoken texts, while the remaining six stem from written texts. In a non-hierarchical analysis, all of these would be independent data points, while in actual fact there are only $n = 3$ spoken texts and $n = 2$ written texts, and the hierarchical analysis that takes this fact into account will be more cautious in estimating the contrast between speech and writing. For further discussions of the potential consequences of applying an ordinary least-squares analysis to nested data, see Hox (2010: 3), Snijders & Bosker (1999: 15–16) and Luke (2004: 6–7). Hierarchical models, then, relax the assumption of independent observations by making dependencies and hierarchies part of the model structure (Hox 2010: 6). Further, missing data (or, in this case, unequal numbers of observations within groups) are unproblematic (Snijders & Bosker 1999: 52, Speelman et al. 2018: 1).

Speelman et al. (2018: 3) suggest that, with regard to random effects, we should consider the levels found in the data to be merely a sample from possible levels. For instance, there may be other genres similar to the ones in ICE, and there are of course more speakers/writers producing language in those genres than the ones that happen to be sampled for our corpora. By specifying a random effect for GENRE, we aim to extrapolate from the genres included in the corpus to the population of genres represented by our sample (cf. Speelman et al. 2018: 3). In contrast, we know the levels or values of our fixed effects. For example, if we use mode of production as a variable, its levels have to be either "spoken" or "written". Combining both random and fixed effects results in a mixed-effects

model. In this study, only GENRE and TEXT are defined as random effects. Thus, results are taken to hold for each of the nine varieties, and varieties are compared to each other, but the study is not designed to make claims that hold for the entire English language complex.[11]

6.3.2 Bayesian statistics

Bayesian data analysis is gaining ground in different empirical disciplines, including linguistics. While it contrasts with classical "frequentist" methods in several important respects, the two approaches will often produce similar results. Bayesian inference is computationally more expensive, both in running the models and in generating meaningful output, e.g. predicted probabilities and the respective visualisations. Its advantages may thus not be immediately obvious, and a few obstacles need to be overcome when adopting it. I will therefore briefly outline some of the properties of Bayesian models and motivate their preference over frequentist ones.[12] Decisions were based both on statistical and practical considerations (cf. Sönning 2020: Section 3.4.2).

1. *Incorporating prior information:* In the Bayesian paradigm, the researcher must incorporate information that is external to the data at hand. Such information is referred to as prior information, or *priors*. A prior is a distribution of likely values established without having seen the data. Another distribution of likely values – the *likelihood* – is generated from the data. Finally, prior distribution and likelihood are fused into a *posterior distribution* of values, the outcome of the analysis. This is Bayes' rule: "the mathematical relation between the prior allocation of credibility and the posterior reallocation of credibility conditional on data" (Kruschke 2015: 99–100; cf. Shikano 2015: 36). Priors can reflect (i) results from previous research, (ii) common-sense assumptions (or "consensual experience"; cf. Kruschke 2015: 27) concerning reasonable outcome values, and (iii) subjective beliefs held by the researcher. Priors can (and should) be criticised if their motivation is not transparent or if they seem to be designed to support a particular hypothesis. If they are carefully motivated and implemented, however,

---

[11]The inclusion of random effects in a model – be it Bayesian or frequentist – results in what is called *shrinkage* (or *pooling*), as discussed by Kruschke (2015: 245–249), Gelman & Hill (2007: 252–259) and Baayen (2008: 275–278). This means that individual estimates in a random-effects structure are partially adjusted towards the general trend (the regression line). However, since it is not discussed or interpreted in this book, I do not discuss this concept in detail here.

[12]See, for example, Kruschke (2015: chapter 11) for a strong argument against classical Null Hypothesis Significance Testing (NHST).

they have considerable advantages. Epistemologically, we should neither blindly trust previous research or common-sense assumptions, nor should we put complete faith in our own data. There is good reason to believe that both components have a contribution to make and should therefore find entry into our analysis. Furthermore, using prior information in regression models also has practical computational advantages, as will be explained below.

2. *Interpretability and statistical thinking:* Compared to frequentist approaches, Bayesian estimation is based on a different notion of probability. Credible intervals in Bayesian analysis (which will be called *uncertainty intervals* in this study) can be directly interpreted as having a certain probability (e.g. 50% or 90%) of containing the model parameter, or the predicted point estimate of the outcome variable. We thus arrive at statements like the following (cf. Figure 7.3 in §7.3): "Based on our statistical model, our prior assumptions and our data, we take the most likely frequency of *though* in HKE to be 72 pmw; there is a probability of 50% that it is found somewhere in the interval of 59–86 pmw, and a 90% probability that it is found within the interval of 44–108 pmw".[13] Direct statements of this kind cannot be made on the basis of frequentist regression models (cf. Bolstad & Curran 2017: 7). However, while this is often adduced as a main advantage of Bayesian statistics, its practical consequences for empirical researchers are in fact rather limited in many cases: Formulating parallel frequentist and Bayesian models for the same dataset will often produce virtually identical results. In this study, considerations listed under the fourth point below played a much more important role.

3. *Analytical efficiency and robustness:* Under certain circumstances, frequentist models may be unable to compute model parameters and fail to converge. This can be due to the complexity of a model, or to a particular data structure (cf. Gelman & Hill 2007: 344). Concerning the first point, a frequentist model has no prior knowledge concerning model parameters. If, therefore, a very large number of parameters (e.g. due to many interaction terms) has to be estimated, all values are considered possible, and the model has to iterate through the entire, multi-dimensional parameter space. Carefully constraining the range of possible parameter values by

---

[13]We also implicitly assume that we have not only considered all relevant predictor variables but also did not make any systematic measurement errors. These conditions are never met completely, and all models will therefore be overoptimistic.

defining priors – even if they are quite weak – can considerably reduce the strain put on such models. Secondly, if there is a lack of variation concerning some parameter (i.e. if certain predictor values correlate perfectly with certain outcomes), frequentist models may fail to converge or will compute unreasonable parameter values (Agresti 2013: 233–235). For instance, if all spoken genres use variant A of a binary variable and all written varieties use variant B, this is likely to result in the non-convergence of a conventional model. While a Bayesian model will still predict very high proportions of variant A for speech and very high proportions of variant B for writing, some (minimal) probability of the alternative variant is generated from the prior, and the model will converge.

4. *Flexible estimates of means, differences and ratios:* Results of Bayesian regression analyses can be summarised and visualised very flexibly, using the R packages selected for this study. The output can be used to calculate point estimates and directly associate them with their statistical uncertainty. For instance, we can not only estimate central tendencies and their dispersions for two conditions, but we can also estimate the difference – be it an absolute difference or a ratio – between the two conditions and its uncertainty, all based on a single model object. This degree of flexibility is beyond current non-Bayesian approaches as implemented using the R package *lme4* (Bates et al. 2020), for instance. Together with the previous point, I regard this as the main practical advantage of Bayesian regression.

In this study, the data were modelled with the utilities in the R package *brms* (Bürkner 2020), which in turn uses Stan (Stan Development Team 2011–2019). The syntax for constructing Bayesian regression models on this basis is rather similar to conventions in non-Bayesian mixed-effects regression modelling with R, e.g. using *lme4* (Bates et al. 2020). With *brms*, it is possible to implement a vast range of model types, including the ones used in this study: negative binomial, binary logistic and multinomial logistic models. The major difference compared to conventional frequentist models is the necessity to specify priors. If the researcher does not do so, *brms* will automatically specify vague priors. However, as explained above, the opportunity to specify priors based on background knowledge and expectations should be actively exploited.

Priors are expected distributions of model parameters, comprised of a central tendency and a dispersion measure. In this study, priors were defined as normal distributions with mean and standard deviation. Priors can be uninformative or informative. Uninformative priors may be entirely "flat", i.e. they make no assumptions about more or less likely parameter values at all, or they may assume

a prior distribution centred at zero and attached to a large standard deviation. Results from Bayesian analyses using such priors will (nearly) coincide with frequentist analyses. Informative priors make assumptions about the most likely parameter value. For example, they might assume that the frequency of a phenomenon is larger (or smaller) in L2 varieties, compared to L1 varieties. Weakly informative priors attach high degrees of uncertainty (i.e. large standard deviations) to such assumptions; strongly informative priors are characterised by a high degree of certainty (i.e. smaller standard deviations). Finally, regularising priors are centred on zero and can again be characterised by smaller or larger standard deviations.[14]

Figure 6.3 illustrates the mechanism whereby a prior distribution of parameter values is confronted with a distribution of values suggested by data (the *likelihood*), and the two are merged into the posterior distribution of the parameter values (the *posterior*). The dotted line in each panel of the figure represents the prior, which is in this case always centred at zero but comes with decreasing standard deviations, from left to right. That is, in panel (d), the researcher is least prepared to accept parameter values that differ from zero, prior to seeing the data. The dashed line in each panel represents the distribution of parameter values that would be considered most likely if only the data were consulted. This distribution is the same in each panel, since our intention is to illustrate the effects of different priors on an analysis. Finally, the solid line represents the posterior distribution of the parameter, which results from combining the prior and the likelihood.
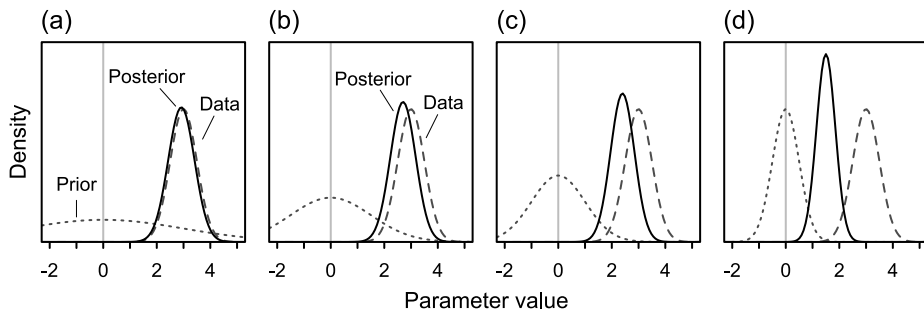
Figure 6.3: Parameter distributions in Bayesian models: priors, data and posteriors

---

[14]Regularising priors are sometimes used as a precaution against overfitting and can be seen as a special case of informative priors. If strongly regularising, they assume a parameter value of zero (i.e. "no effect") in combination with a small standard deviation. It is then relatively difficult for the data to prevail over such priors.

We can see that in Figure 6.3a, the prior essentially lets the data speak for themselves; in panel (b), the posterior is pulled slightly towards zero; in panel (c), this effect is more pronounced; and in panel (d), prior and likelihood are equally informative, so that the position of the posterior is intermediate between them. Many of the priors used in this study resemble the one used in Figure 6.3b, with a central tendency at zero and a relatively generous standard deviation. Such priors, while they will easily yield to strong data, will nevertheless provide computational support to complex models by placing gentle constraints on the possible parameter space. In some cases, priors are selected so as to reflect the direction of an effect in previous publications, but they, too, come with a large standard deviation.

Running a Bayesian model based on priors and data returns sampled combinations of parameter values (the posterior sample), and each parameter can then be described concerning its central tendency and dispersion. Sampling happens via a method called Markov Chain Monte Carlo, or MCMC (Agresti 2013: 23, Kruschke 2015: 115–116, 144, Shikano 2015: 37). McElreath (2020: 263–298) provides an accessible discussion of the procedure. Essentially, a particular, randomly selected combination of parameter values is likely to make it into the posterior sample if it describes the data relatively well. Therefore, a large number of likely combinations of parameter values (positioned in regions of relatively high probability densities) and a correspondingly lower number of *un*likely combinations are returned by the sampling process.

The MCMC routine is partitioned into several parallel processes (or chains) that should be generally aligned, or correlated, and each chain contains a specified number of warmup samples – used to calibrate the sampler – that are excluded from the final sample (Gelman & Hill 2007: 356). For example, if there are $n = 3$ chains with $n = 4,000$ iterations each, and if there is a warmup phase of $n = 500$ iterations, the total number of data points in the posterior sample will be $3 \times (4,000 - 500) = 10,500$ samples, each of which contains a unique combination of model parameter values.

6.3.3 Types of regression used in this study

The models used in this book all qualify as generalised linear models, since the outcome quantity is modelled on a transformed, nonlinear scale (Raudenbush & Bryk 2002: 291). This transformation of original quantities (here: rates of occurrence and proportions) happens via *link functions*, as will be explained below.[15]

---

[15]Strictly speaking, linear models are a special case of generalised models (see Gelman & Hill 2007: 109) that use the *identity function* as a link between the outcome and the statistical model, i.e. the outcome is left untransformed during the estimation process.

6.3.3.1 Negative binomial count regression

Since rates of occurrence (i.e. how often something happens) have a lower bound at zero, they are usually analysed using count regression models; ordinary regression models are not recommended for such outcomes (Long 1997: 217–218). Several types of count models are available, e.g. the Poisson model and the negative binomial model. For reasons discussed by Agresti (2013: 7, 127), Long (1997: 230) and Gelman & Hill (2007: 115–116), I will use the negative binomial model in this book (cf. Cameron & Trivedi 2013: 1). Binomial models address certain issues in Poisson regression models, for instance the problem of overdispersion: Poisson regression assumes that predicted means have a constant variance which equals the mean, when in reality means become increasingly overdispersed as they grow larger, i.e. the variance exceeds the mean, which has undesired effects on the performance of the model. The four equations in (80) show the steps involved in modelling frequency of occurrence using count regression models.

(80)   Count regression: outcome (a), log link (b), model (c), exponential (d)

   a. $y = \dfrac{n}{N} = R$ c. $\eta = a + xb$

   b. $\eta = \ln(R)$ d. $y = \exp(\eta)$

Ultimately, we are interested in the outcome y, which is the rate of occurrence R of a phenomenon, calculated by dividing the raw number of occurrences n by some baseline N, e.g. the number of words in a corpus. For count models, the log link function shown in part (b) of the equation is used, i.e. rates are transformed using the natural logarithm (Cameron & Trivedi 2013: 36, Kruschke 2015: 705–706, Molenberghs & Verbeke 2005: 32, Agresti 2013: 115). This results in values of $-\infty < \eta \leq 0$. Logged rates are then modelled as shown in (80c): The average value (or intercept) a and the predictor coefficient b are estimated in such a way that, in combination with different predictor values (like x), they provide a good approximation of the outcome η.[16] Estimated values directly based on the values of coefficients in the posterior are on this logged scale and need to be retransformed as shown in (80d), obtaining values of $0 < R \leq 1$ (Agresti 2013: 125). To arrive at the typical representation of rate per million words, we finally need to multiply by 1,000,000, a step that is not shown in (80).

---

[16]Note that, unlike the analyses in this book, the schematic example used here is very simple and non-hierarchical.

6.3.3.2 Binary logistic regression

If an ordinary regression model is applied to binary outcomes, nonsensical estimates of $p > 1$ or $p < 0$ may result (Snijders & Bosker 1999: 211, Luke 2004: 53; cf. Best & Wolf 2015). Partly for this reason, regression models with binary response variables, like count models, depend on a nonlinear link function. In analogy to (80) above, the equations in (81) show the steps involved in a binary logistic regression analysis.

(81)   Binary logistic regression: outcome (a), logit link (b), model (c), logistic (d)

    a. $y = \dfrac{n}{N} = p$ c. $\eta = a + xb$

    b. $\eta = \ln\left(\dfrac{p}{1-p}\right)$ d. $y = \dfrac{\exp(\eta)}{1 + \exp(\eta)}$

The outcome y is a proportion p, which is the fraction of the total number of times a variant occurs over the total number of contexts in which it *could* have occurred; it takes values of $0 < p < 1$. In binary logistic regression, the logit link function shown in (81b) is used to transform an expected proportion to log odds (or logits), i.e. the natural logarithm of the odds (Snijders & Bosker 1999: 212, Luke 2004: 53, Agresti 2013: 115, Gelman & Hill 2007: 80, Kruschke 2015: 622). For instance, the odds for a proportion of 0.8 are 4:1 (or simply 4), and the corresponding log-odds (or logit) value is $\ln(4) = 1.39$. For the inverse case with a proportion of 0.2, the odds are 1:4 (or 0.25), yielding a logit of $\ln(0.25) = -1.39$. Complementary logits (e.g. for pairs of proportions like 0.8 and 0.2, or 0.45 and 0.55) are symmetrically arranged around zero, with the same unsigned values, and they lie in the range of $-\infty < \eta < +\infty$. Logits can then be modelled linearly, as shown in (81c). For the communication of results, it is desirable (and perhaps necessary) to transform estimated values back into proportions, using the logistic function illustrated in formula (d) above (Luke 2004: 56). To obtain percentages instead of proportions, values thus obtained simply need to be multiplied by 100. In a binary regression model, predictor coefficients refer to a difference in the probability (expressed as log odds) of observing one of the two outcome variants relative to the other. In such models, it is usually enough to show results for the "target" category, because we can infer the relative frequency of the reference category: $p_1 = 0.8$ corresponds to $p_0 = 0.2$, for instance, and vice versa. However, in order to estimate values for the baseline category, we simply need to insert the value 1 for the numerator in formula (d).

6.3.3.3 Multinomial logistic regression

If there are more than two possible outcome categories, multinomial regression may be used. To this end, one level of the categorical response serves as a baseline category. Multinomial regression can be thought of as a series of binary models, since the odds of each of the remaining levels are compared to the odds of the baseline category (Agresti 2013: 293–294). The link function is once again the logit link, but in this case, there are logits for all categories except the baseline. In essence, however, multinomial regression can be regarded as an extension of binary logistic regression (Gelman & Hill 2007: 119), or as Long (2015: 173) puts it, they are "sets of binary regressions that are estimated simultaneously". In practice, this means that the specification of priors – if a Bayesian approach is used – can be more complex, as different specifications can be made for different outcome categories. More importantly, the re-transformation of logits into predicted probabilities is also more involved (cf. Agresti 2013: 296–297). If we assume three outcome categories (e.g. in the selection of one of the three concessive conjunctions in this book; cf. Chapter 10), the equations in (81a–c) above can be directly applied to multinomial regression, except that the number of parameters is multiplied by the factor $c-1$, c being the total number of possible outcomes. With three outcome categories, that is, y, η, a and b remain unspecified for the reference category, but for the other two categories y_1, y_2, η_1, η_2, a_1, a_2, b_1 and b_2 are required. Accordingly, the overall number of model parameters will be higher, compared to binary logistic regression. In (82), the logistic function for transforming summed logits back into proportions is shown for multinomial models with three outcome categories (see Long 2015: 176). In this scenario, the index i takes two values, representing the alternative (non-baseline) categories.

(82)   Logistic function for tricategorical multinomial response models

$$y_i = \frac{\exp(\eta_i)}{1 + \exp(\eta_1) + \exp(\eta_2)}$$

Summed logits for different outcome categories are inserted into the equation, where they are exponentiated to obtain odds, then transformed to obtain proportions. The denominator is fixed, i.e. it always consists of the summed odds of all categories, including the reference category, which is represented by the value 1. So it is only the numerator that varies, depending on which category is being estimated. If the baseline category is of interest, the numerator will take the value 1 (see the discussion of binary logistic regression above).

6.3.4 Comments on random effects and centred predictors

One property of logits is that they take increasingly extreme values for underlying proportions close to $p = 0$ and $p = 1$, respectively. In random-effects models with a nonlinear link function, this may have effects on the predicted probabilities that run counter to our intuitions, and to what we see in the raw data. In Table 6.1, let us assume that we compare proportions of some categorical outcome category (e.g. *even though*) in five texts. Proportions 1, 2 and 3 are very close to zero, proportion 4 is substantially higher (corresponding to 4.3%), and proportion 5 is higher still (corresponding to 25.1%). While the mean proportion is 0.063, a much lower mean results if we first average the logits corresponding to proportions and then reconvert that average into a proportion. Instead of 0.063 (i.e. 6.3%) we obtain only 0.019 (i.e. 1.9%).

Table 6.1: The distorting effect of averaging logits

	Text					M	
	1	2	3	4	5		
p	0.003	0.005	0.012	0.043	0.251	0.063	
LOGIT	−5.81	−5.29	−4.41	−3.10	−1.09	−3.94	$\rightarrow p = 0.019$

That is, predicted probabilities in random-effects models will be biased towards the extremes (0 and 1) if there is a nonlinear link function, and if the variance between groups is large – usually because some groups (e.g. texts/speakers) behave near-categorically (see Molenberghs & Verbeke 2005: 299, Agresti 2013: 495–498). Even if there are groups with considerably higher proportions of a variant, their effect on the overall outcome will be overruled by groups with very low proportions, because random intercepts (and slopes) are centred around the mean logit. Thus, there may be a marked contrast between predicted probabilities and what we see in the raw data.[17]

As a strategy to cope with this phenomenon, Molenberghs & Verbeke (2005: 301) suggest that conditional means (i.e. predicted percentages or proportions) should be established for different levels of a random variable, and that these should then be averaged. As a practical solution, they further suggest that for each random coefficient, a large number of values should be randomly sampled

---

[17]Of course, model-based estimates will always differ more or less dramatically from what the raw data tell us, but one's suspicions should be raised if there is a fundamental difference, not one of degree.

from the distribution returned by the model (which is always a standard deviation attached to a mean of zero). For each randomly sampled value, the conditional mean is calculated (as a proportion, or percentage), then the average of these values is established. In effect, we estimate the outcome (proportions/percentages) for a large number of (fictive) units and then show the average unit; by contrast, in the easier, default approach, we average values whilst still operating within the nonlinear link function. Since in a Bayesian approach the estimated standard deviation of a random effect will be different in each line of the posterior sample, we need to conduct this random-sampling approach for each line. The procedure is unproblematic, but computationally expensive.

A similar issue arises in the fixed part of binary logistic or multinomial models with centred predictor variables. Take, for instance, the predictor SPOKEN.CT, which in this study takes values of +1 ("spoken") and −1 ("written"). If the average proportion of an outcome is 0.05 in writing and 0.35 in speech, the estimated average proportion that should result if we constrain the predictor to take its centred (or neutral) value of zero is $p = 0.20$, as shown in Table 6.2.

Table 6.2: Skew introduced by centred predictors

	written	spoken	M	
p	0.05	0.35	0.20	
LOGIT	−2.94	−0.62	−1.78	$\rightarrow p = 0.14$

However, since the model operates on the logit scale and since we use a centred predictor, the mean value that we obtain on the basis of SPOKEN.CT will introduce a skew. The value in the example is −1.78 and yields a proportion of $p = 0.14$ when reconverted. That is, average values in regression models with nonlinear link functions tend to be biased towards the extremes ($p = 0$; $p = 1$) if we rely on the centred (zero) value of a predictor. The strategy adopted in this book to counteract this effect was to make concrete posterior predictions on the percentage scale for all specific conditions, and then to average across them. The technical details can be traced in the published analysis scripts (see §1.4).

6.3.5 Estimation and visualisation

Each item in the posterior sample, then, consists of one complete set of values for all model parameters. In the estimation process, we calculate one outcome percentage for each line of the posterior sample, based on the predictor values

for the condition of interest. It is also possible to calculate values for two conditions and directly subtract one from the other – again this happens line by line, yielding a distribution of possible values. We can summarise these values in various ways, e.g. by reporting their means and standard deviations, or their medians along with certain quantiles. These values can then be plotted and interpreted: We can make a statement about the most likely, central outcome value, and we frame intervals that contain the central 50% or 90% of values, for example. In this study, these are called *uncertainty intervals* and can be regarded as analogues to frequentist *confidence intervals* (Agresti 2013: 23), but see Gelman & Greenland (2019) for a terminological discussion. Uncertainty intervals can also be calculated for individual parameters (coefficients) if we are interested in their precision. In the text, a 90% uncertainty interval will sometimes be stated along with a point estimate (i.e. the "best guess" for a value of interest). Thus, in §10.2.2, the proportion of *even though* in anticausal CCs is given as "54.5% [43.9; 64.9]" for CanE, which indicates that in this variety the median estimate is 54.5%, with a 90% uncertainty interval extending from 43.9% to 64.9%.

Figure 6.4 illustrates what conditional effects plots look like in this book. The point estimate shown as a filled black circle at the centre of the distribution represents the median posterior predicted probability, i.e. the median of all values sampled from the posterior for a given condition – this would be the most typical outcome value. The two intervals extending upwards and downwards from the median represent 50% and 90% uncertainty intervals: The thicker bar covers values between the 25th and 75th percentiles and the thinner bar extends to the 5th and 95th percentiles of the posterior predicted probability of the outcome.
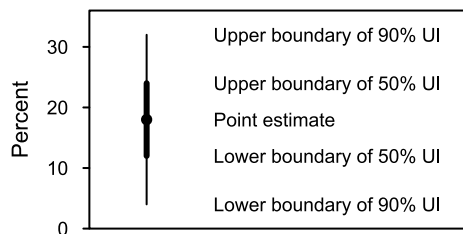
Figure 6.4: Typical plotting symbols used in the visualisation of results

In all analyses in this book, outcomes and effects will be expressed in readily understood quantities, i.e. rates of occurrence and percentages of variant outcomes, the latter being based on predicted probabilities (e.g. Long 2015: 173). In the comparison of conditions, differences and ratios will be used. I follow Best & Wolf (2015), who advise against using logits or odds ratios when interpreting

logistic regression models, since, in isolation, they show little more than the direction (and general strength) of an effect. Odds ratios may be easier to interpret than logits, but they, too, do not make the magnitude of an actual change in probabilities between conditions transparent (Long 2015: 188). Presentations of results in this study will thus consist of back-transformed outcomes that are appropriately visualised (cf. Bolstad & Curran 2017: 31), using the packages *lattice* (Sarkar 2008, 2021) and *latticeExtra* (Sarkar & Andrews 2019). Regression tables are relegated to the online appendix as they contribute little to our understanding of actual patterns (cf. Gelman & Hill 2007: 457). Rather than *p*-value significance testing of individual coefficients, more informative estimation approaches are used (Gelman & Hill 2007: 22–23, Cumming & Calin-Jageman 2017).

6.3.6 Specification of variables and model selection

According to Baayen (2008: 241–242), the distinction between fixed and random effects is not primarily about grouping structures, but about "repeatable levels" (i.e. predictor levels/values that are known) and factors "with levels randomly sampled from a much larger population". Mixed-effects models include both types of variables. Table 6.3 lists the variables used in this book, ordered according to the three types "outcome", "predictor" (i.e. fixed-effect), "control" and "random". It also specifies which variables are relevant in which chapter(s), as well as the values each of them can take.

The variable LENGTH.CT is labelled as a control variable, since no research interest is attached to it and it is not interpreted in detail. Actual values for this variable were in the range of 1–36 words. They were logged and centred round the geometric mean of all tokens at 8.8 words. Thus, the shortest subordinate clause (length = 1 word) and the longest subordinate clause (length = 36 words) translate into values of LENGTH.CT = $\ln(1) - \ln(8.8) = -2.17$ and LENGTH.CT = $\ln(36) - \ln(8.8) = 1.41$, respectively. Possible values for all other variables are more straightforward, as documented in Table 6.3. For the five models that were used, Table 6.4 provides a concise definition of the outcome variables, states the labels given to the models, and makes reference to the chapters and appendices relevant with regard to each model.

Agresti (2013: 210) discusses model selection strategies that build more complex models up from more basic ones ("forward") and strategies that start from maximally complex models and reduce them in a stepwise fashion ("backward"). He concludes that both approaches can be problematic and may yield suboptimal models, and that selection procedures for variables should be applied with due caution:

Table 6.3: Specification of variables; for uncentred categorical variables, baseline categories are stated first, predicted categories and non-baseline predictor levels appear in bold print.

Variable	7	8	9	10	11	Values/levels
Outcome						
COUNT	•	•				logged frequency (pmw): $\ln(f)$
FINAL			•			nonfinal, **final**
MARKER				•		*although*, ***though***, ***even though***
NONFIN					•	finite, **nonfinite**
Predictor						
SPOKEN.CT	•	•	•	•	•	written = −1; spoken = +1
MARKER	•				•	*although*, ***though***, ***even though***
TYPE		•				dialogic, **anticausal**, **narrow-scope**, **epistemic**
ANTI.CT			•	•	•	dialogic = −1; anticausal = +1
FINAL.CT				•	•	nonfinal = −1; final = +1
Control						
LENGTH.CT			•			continuous; range = [−2.17; 1.41]; $M = 0$
Random						
GENRE	•	•	•	•	•	30−32 levels per variety
TEXT			•	•	•	177−262 levels per variety

Table 6.4: Model designations

Outcome	Label	Ch.	Details
Rate of occurrence: Markers	Model A	7	Appendix B.1
Rate of occurrence: Semantic types	Model B	8	Appendix B.2
Percentages: Clause positions	Model C	9	Appendix B.3
Percentages: Markers	Model D	10	Appendix B.4
Percentages: Clause types	Model E	11	Appendix B.5

> [S]tatistical significance should not be the sole criterion for inclusion of a term in a model [...]. It is sensible to include a variable that is central to the purposes of the study and report its estimated effect even if it is not statistically significant. Keeping it in the model may help reduce bias in estimated effects of other predictors and may make it possible to compare results with other studies where the effect is significant, perhaps because of a larger sample size. Algorithmic selection procedures are no substitute for careful thought in guiding the formulation of models.

Winter (2020: 276–279) provides excellent arguments along similar lines, and the present study follows this kind of thinking: A carefully selected set of variables is included and kept in the model even if one or several of them have only a small effect and/or come with a high degree of uncertainty and would therefore be non-significant in a frequentist model (see also Tizón-Couto & Lorenz 2021). Thus it is ensured that the link between the theoretical background of the study and the results remains unbroken at all times, even if this means that the presented models are usually not the most parsimonious ones. In order to enable comparison in the present study, it is crucial that models should have the same structure across varieties.

7 Frequencies of conjunctions

This chapter follows up on Hilpert's (2013b: 462) suggestion that frequencies of occurrence can be relevant for studying constructions and constructional variation and change. However, the purely frequency-based approach, common in much of traditional corpus linguistics, needs to be viewed critically: Frequencies of forms result from the need to express certain semantic or grammatical relations, and questions concerning those underlying factors should therefore be primary. For example, higher or lower frequencies of particular concessive conjunctions in certain varieties may be the result of (i) differences in the proportion of concessives that are expressed by means of subordination, (ii) general differences in the frequency of concessives (that is, subordinating and other), (iii) a topic bias in the sampled corpus material that favours or disfavours the use of concessives, or (iv) a combination of several of these factors. This chapter nevertheless pursues the traditional approach of counting surface frequencies of conjunctions. In combination with Chapter 8, it serves as a point of departure for later analyses.

Concerning the text frequencies of conjunctions, no precise hypotheses were formulated in §5.3, mainly for the reason that more profound insights are expected from a variationist perspective, i.e. when investigating the relative frequencies (proportions or percentages) of *although*, *though* and *even though* as variant forms. However, the general expectations would be that (i) the frequency ranking found in the literature will be confirmed, namely *although > though > even though*, perhaps with some uncertainty as to the position of the latter two, and that (ii) all three conjunctions should be more frequent in writing, perhaps less so with regard to *even though*. No general difference between L1 and L2 varieties is anticipated. Based on some of my earlier research, §7.1 briefly inspects the frequencies of a range of markers (conjunctions, prepositions/adpositions and conjuncts) beyond the ones targeted in this monograph. Section 7.2 introduces the statistical model used for the main frequency analysis presented in §7.3, which then focuses on the three conjunctions *although*, *though* and *even though*. Results are summarised in §7.4.

7.1 Overview: Frequencies of different concessive markers

This section provides some minimal context for the three conjunctions under investigation in this study by showing their rates of use relative to other, functionally related markers. A general overview of the frequencies of concessive markers that belong to different grammatical categories is given in Figure 7.1, which is based on data from twelve ICE corpora and thus goes beyond the selection of varieties included in the present volume (see Schützler 2018b). Panel (a) shows global mean frequencies while panel (b) contrasts frequencies in speech and writing.[1]

Three broad groups can be identified. The conjunct *however* and the subordinator *although* are the most frequent connectives; *though* (both as a conjunction and a conjunct), *despite* and *even though* constitute a cluster of markers that are of intermediate frequency; a third group of relatively rare markers comprises *nevertheless*, *in spite of*, the conjunction *however* and the adposition *notwithstanding*. For patterns in individual varieties that diverge slightly from this general picture, see Schützler (2018b: 158).

Figure 7.1b indicates that in the vast majority of cases concessive markers are considerably more frequent in written language. The only items that deviate from this pattern are the conjunct *though*, which is preferred in speech (cf. Schützler 2020b), and the conjunction *even though*, which seems to be equally frequent in both modes of production. The result for the latter anticipates tendencies also found in the more detailed, regression-based analyses in §7.3 below, with more insights provided in subsequent chapters. Counts of the type presented in Figure 7.1 can go some way towards an estimate of the total number of concessives used in varieties of English. However, one must bear in mind that it is also possible for concessive meaning to be expressed without an explicit grammatical marker, and that the coordinator *but* poses a certain problem since it is very frequent and quite variable in function (e.g. concessive vs purely adversative).

---

[1]Additional varieties included for Figure 7.1 were (i) US-American English, represented by a combination of ICE-USA and the *Santa Barbara Corpus of Spoken American English* (Du Bois et al. 2000–2005); (ii) Philippine English; and (iii) New Zealand English. The procedure for both parts of the figure was to determine for each marker in each variety frequencies in speech and writing, as well as their geometric mean (cf. Footnote 2 on p. 119). Each data point shown in the figure was then arrived at by calculating the geometric mean of the twelve respective variety-based values (cf. the online appendix at https://osf.io/m4tfc/). Frequencies were thus not estimated using a regression model (cf. §7.3) but simply counted and normalised within the two broad categories of speech and writing.

(a) Global means

(b) Speech vs writing
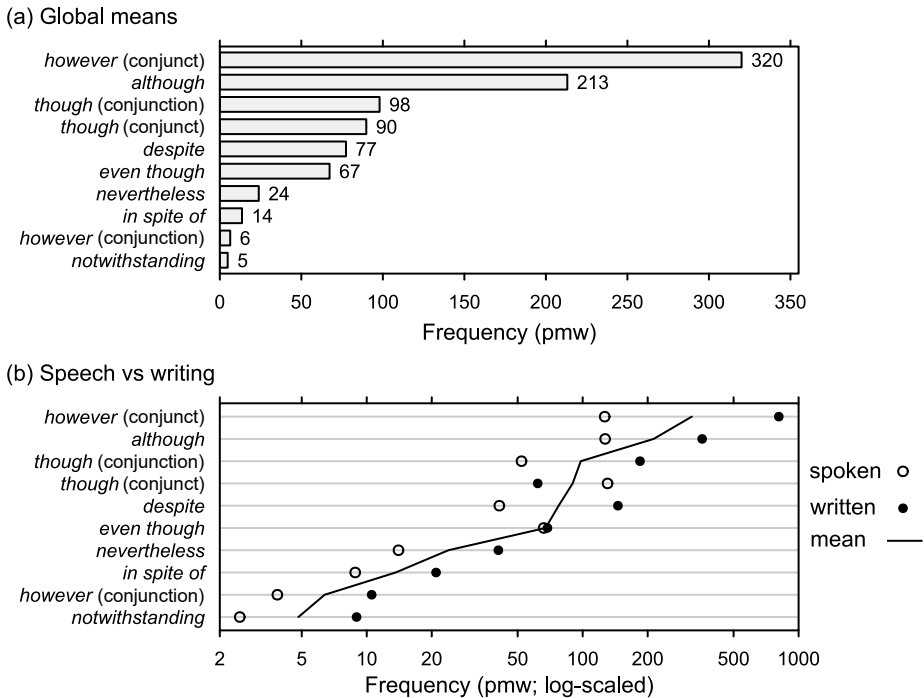
Figure 7.1: Frequencies of different concessive markers (ICE)

7.2 Statistical model

Frequencies of all three conjunctions are estimated with a Bayesian negative binomial mixed-effects regression model, which is given the denomination "Model A" and breaks down into nine variety-specific submodels of exactly the same form, each based on the respective subset of the data. The output from these models is used in all analyses in §7.3. The only cluster variable in the model is GENRE: The smallest unit of observation in the (negative binomial) count model is the individual text, which is why TEXT does not enter the model as a cluster variable. As discussed earlier, variety, too, does not feature as a cluster variable, due to the one-model-per-variety approach that was taken – a characteristic shared by all models in this study. There are only two predictor variables in the fixed part of Model A, (i) mode of production and (ii) the marker itself (*although*, *though*, *even though*), corresponding to the independent variables SPOKEN.CT and MARKER. The model syntax is shown in (83); for a definition of variables, see Table 6.3 in §6.3.6. The predictors SPOKEN.CT and MARKER interact in the model. Additionally, slopes for MARKER are specified as varying randomly across GENRE.

The component labelled "offset" refers the model to the variable LOG_WORDS, which quantifies the logged number of words per text. Note that this variable is not specified in Table 6.3.

(83)   Model A: Syntax

```
count ~ spoken.ct * marker
        + (marker | genre)
        + offset(log_words)
```

Appendix B.1 provides more information concerning token numbers, the number of levels of the random factor GENRE, the implemented priors and the number of posterior samples. Data, scripts and comprehensive regression tables are published online (cf. §1.4).

7.3 Results

There will first be a discussion of global frequency patterns, i.e. a comparison of the estimated average frequencies of the three markers in all varieties, which is then unfolded into a comparison of frequencies in speech and writing. Finally, the focus shifts from a variety-based approach to a marker-based approach by ranking for each connective the conditions that favour (or disfavour) its occurrence. This latter part does not provide new information but a new perspective.

A first general assessment is shown in Figure 7.2. Here, median frequency estimates of the three conjunctions are summed for each variety, and varieties are ranked by this total number of CCs. The four L1 varieties are shown in black, while the five L2 varieties are shown in grey.

This basic display highlights that there is a core frequency range for subordinating CCs, which extends from 334 pmw (in JamE) to 388 pmw (in SingE), suggesting a fairly stable rate of use of this kind of construction in most varieties. However, the total number of CCs is substantially higher in BrE (537 pmw) and substantially lower in NigE (284 pmw). Figure 7.2 draws attention to this phenomenon, which tends to be obscured in more detailed visualisations (e.g. in Figure 7.3 below), and raises the question of its potential implications. It seems difficult to motivate the dramatic difference between NigE and BrE. Does it arise (i) because speakers and writers of these varieties stand out in using (or not using) CCs as a semantico-pragmatic device, (ii) because these particular components of ICE contain texts (i.e. speakers/writers) that are atypical, from a cross-varietal perspective, or (iii) because the subject matters in the respective corpora happen
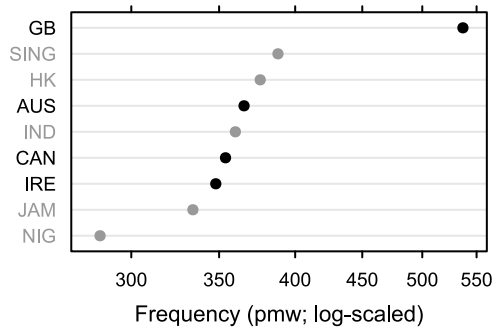
Figure 7.2: Summed frequencies of concessive subordinators in nine varieties; black = L1 varieties, grey = L2 varieties

to favour (or disfavour) the use of CCs? More speculative reasons of this kind could be adduced, but it should be clear that purely frequency-based results of this kind are difficult to interpret. Similar concerns apply when inspecting the log-scaled visualisations further below: Since intra-constructional semantics and the arrangement of subordinate and matrix clauses in a CC are hypothesised (and indeed shown) to play a role in the choice of conjunction (see Chapter 10), a frequency analysis that blinds itself to these factors must be of limited explanatory power. Thus, it cannot be emphasised enough that the present chapter should be understood as a general background against which the results in later chapters emerge all the more clearly in their interpretability. At the same time, it represents a traditional corpus-linguistic approach based on surface frequencies, which makes it comparable to earlier studies.

Frequencies of the three conjunctions *although*, *though* and *even though* in all nine varieties are plotted in Figure 7.3, based on the respective component models of Model A (for a concise summary of model parameters, see the online appendix). Effects of the mode of production are controlled for by estimating average values based on the written and spoken conditions.

As expected, *although* is the most frequent one of the three conjunctions in all varieties except IndE, with an average text frequency of 184 pmw across varieties (not shown).[2] In six out of nine varieties, *though* is the second most frequent conjunction, with an average text frequency of 103 pmw. Finally, *even though* is

---

[2]For the calculation of this and the following two values, the geometric mean was used, i.e. the average frequencies (pmw) in all nine varieties were logged, averaged, and then re-transformed by exponentiation. Thus, the result matches the visual impression conveyed by the plot. The following R function was used, where x is a vector of values to be averaged using the geometric mean: MGeom = function(x) {exp(mean(log(x)))}.
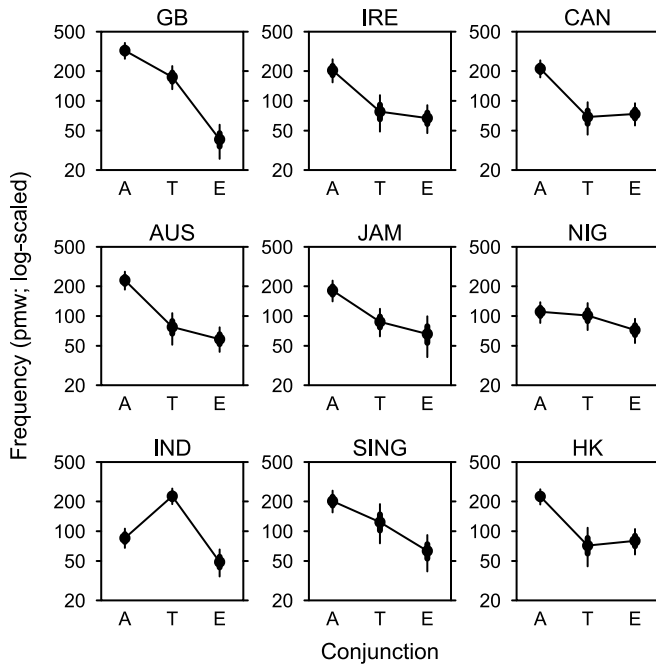
Figure 7.3: Average frequencies of concessive subordinators; A = *although*, T = *though*, E = *even though*

generally least frequent, with only two exceptions (CanE and HKE); its average text frequency is 62 pmw. These findings are in reasonable agreement with what is shown in Figure 7.1 above.

Against the background of general differences between varieties outlined above, the following paragraphs will focus on frequency differences between spoken and written language in the nine varieties. Based on the same statistical model (Model A), the approach here is not to control for mode of production but to show two estimates for each of the three conjunctions in each variety. The visualisation in Figure 7.4 is subdivided into nine parts, corresponding to varieties, each of which takes two perspectives. In the lower panel, estimated text frequencies are plotted, showing expected values in speech and writing. Values for the three conjunctions are connected with dotted lines in each mode of production to facilitate the comparison of frequency patterns (cf. Schützler 2023). In the upper panel of each subplot, differences between frequencies in the two modes are highlighted by focusing on the frequency ratio of f_W divided by f_S. Like the frequency values themselves, this measure of difference is log-scaled to ensure that the relative differences come across more clearly in the visual display (cf. Schützler 2023).

In the vast majority of cases, the three markers are more frequent in writing, which is readily seen when inspecting the upper panels for the nine varieties in Figure 7.4: Virtually all median ratios are greater than (or equal to) 1, with a single exception in NigE (see below). The most extreme frequency difference between writing and speech is found for the conjunction *though* in HKE, with $Ro_{W/S}$ = 6.7 [2.9; 16.2]. For *even though*, the difference between modes is not substantially different from Ro = 1 in seven out of the nine varieties, namely SingE (1.7 [0.8; 3.6]), IndE (1.6 [0.9; 2.9]), IrE (1.4 [0.8; 2.8]), CanE (1.3 [0.8; 2.1]), JamE (1.2 [0.5; 2.7]), AusE (1.0 [0.6; 1.8]), and NigE (0.9 [0.5; 1.5]). In NigE, the ratio is even slightly in favour of spoken language. In BrE and HKE, the written-to-spoken ratio for *even though* is most substantially different from 1, with $Ro_{W/S}$ = 2.1 [1.0; 4.4] and $Ro_{W/S}$ = 1.9 [1.1; 3.4], respectively. The general difference between *although* and *though* (treated as a pair) on the one hand and *even though* on the other is highlighted by several "hockey-stick" patterns in the upper panels of Figure 7.4.

Using frequency differences between speech and writing as a very rough indicator of stylistic function, it appears even from the purely visual inspection of Figure 7.4 that both *although* and *though* are more sensitive (or specialised) in this regard, being much more common in writing than in speech; *even though*, on the other hand, is much more evenly distributed between modes of production. If we average across the written-to-spoken ratios of all varieties for the three conjunctions, using the geometric mean (cf. Footnote 2 on p. 119), this impression is fully confirmed: The average written-to-spoken ratio is 3.0 for *although*, 3.6 for *though*, but only 1.4 for *even though*.

The effect of mode has the same direction in L1 and L2 varieties, but in the latter group it seems to be smaller for the two more frequent conjunctions, with a (geometric) mean ratio of 2.8 for *although* (as compared to 3.1 in L1 varieties) and 2.9 for *though* (as compared to 3.8 in L1 varieties). For *even though*, the two values are about the same (1.3 in L2; 1.4 in L1). If we again accept speech and writing as very rough stylistic categories, the more level pattern in L2 varieties at least for *although* and *though* can very tentatively be interpreted as a lack of differentiation according to Schneider's (2003) Dynamic Model (cf. §4.3.2). As pointed out above, however, conclusions of this kind must be tentative if drawn on the basis of a model that is not truly multifactorial as it ignores other underlying functional and formal factors of potential importance. It will therefore be necessary to revisit the results presented here when discussing findings in Chapter 10.

Further interesting nuances are revealed as we shift our perspective by focusing on speech and writing in isolation, taking a step back from the direct comparison of the two modes of production. In writing, IndE constitutes the single exception to the otherwise perfectly regular ranking $f_A > f_T > f_E$, with *though*
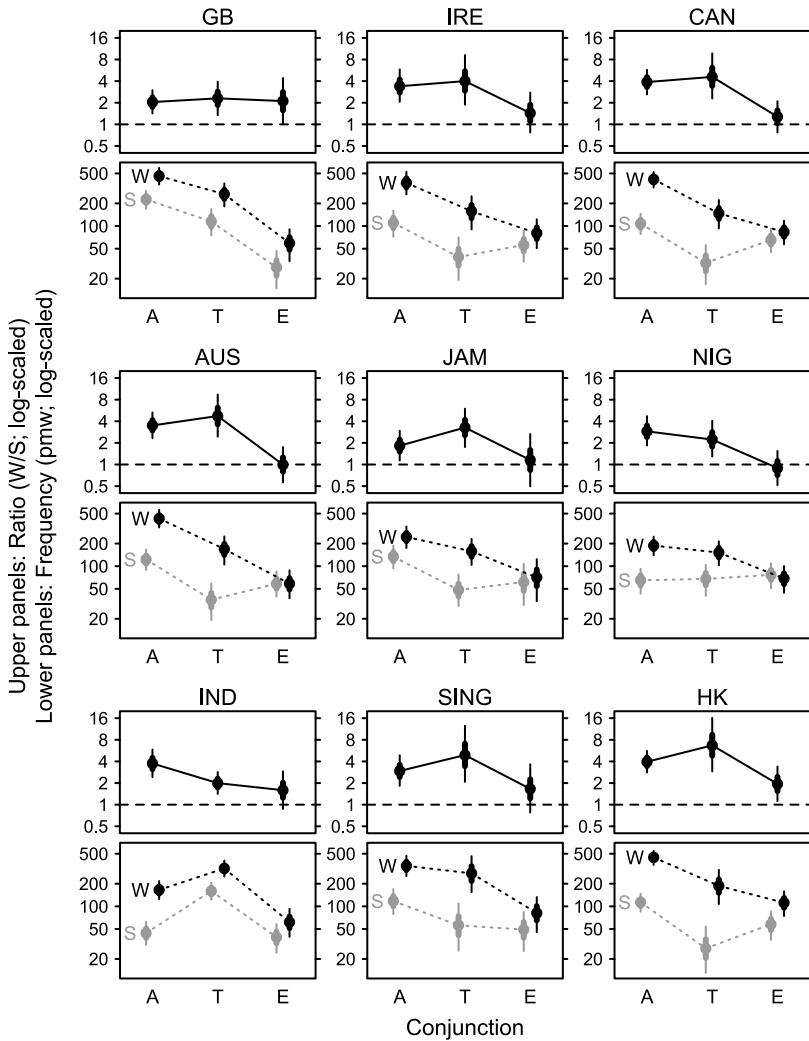
Figure 7.4: Frequencies of concessive subordinators in speech and writing; A = *although*, T = *though*, E = *even though*, W = written, S = spoken

being the most frequent conjunction of the three. The regularity of the pattern in all other written varieties suggests that there is considerable agreement as to which conjunctions are most generally usable in this mode, and this is consistent with McArthur's (1987) postulation of a written World Standard English (cf. Figure 4.4b on p. 69). In speech, the frequency ranking of the three conjunctions is much more variable. Four different patterns exist, the most common one being $f_A > f_E > f_T$, which is found in IrE, CanE, AusE, JamE and HKE, followed by the pattern $f_A > f_T > f_E$ in BrE and SingE. IndE has the unique pattern $f_T > f_A > f_E$, and NigE also stands out in having very similar (low) frequencies of both *though* and *even though*, which can strictly be ranked as $f_E > f_T > f_A$. Thus, four out of the six possible frequency rankings do in fact occur in spoken varieties, while patterns in writing are essentially uniform. From a general perspective, the conjunction *even though* seems to be characterised by a considerably higher (relative) rate of occurrence in spoken English, the result at the level of the individual variety very often being a pronounced difference in pattern between speech and writing – only BrE, SingE and IndE are characterised by the same general ranking of conjunctions in both modes.

Figure 7.5 focuses entirely on the conjunction *although* and ranks all spoken and written varieties (*n* = 18) by their absolute (normalised) text frequencies of this marker. This display and its minimal discussion does not go beyond Figure 7.4 above, but it arranges the information in a different way and thus provides another perspective on the data. On the one hand, the actual frequency range of *although* [44; 461] is more clearly visible here. On the other hand, the black and white boxes on the right highlight the ranking of conditions according to variety type (L1 vs L2) and mode of production. This part of the figure is further supported by triangular indicators that show the mean ranks for the two groups that are compared in each column, using the respective colours. We see that *although* is more frequent in L1 varieties (mean rank: 8.0) than in L2 varieties (mean rank: 10.7), but the more striking contrast is between written and spoken varieties, with mean ranks of 5.2 and 13.8, respectively.

The same perspective is taken for the conjunction *though* in Figure 7.6. Compared to *although*, the range of median values is shifted towards lower values [28; 319].[3] In terms of absolute text frequencies, this marker is somewhat more common in the L2 varieties under investigation (mean rank: 8.6) compared to the L1 varieties (mean rank: 10.6). Once again, frequencies in written varieties are much higher than in spoken varieties, with mean ranks of 5.4 and 13.6, respectively.

---

[3]The range for *though* also appears to be narrower in absolute terms, but on a logarithmic scale the difference between maximum and minimum is very similar for both conjunctions.
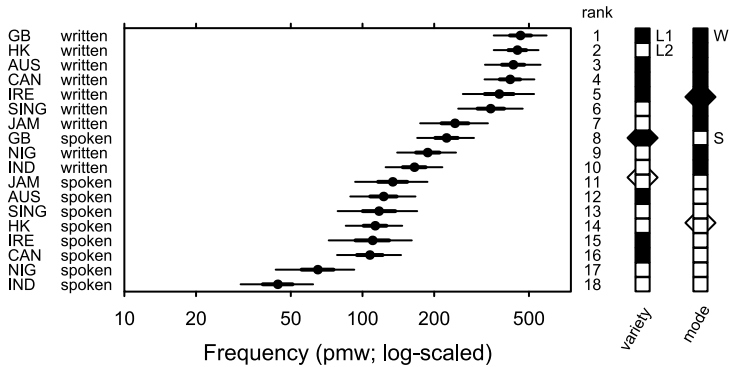
Figure 7.5: Text frequencies: Ranking of specific conditions for *although*; W = written, S =spoken
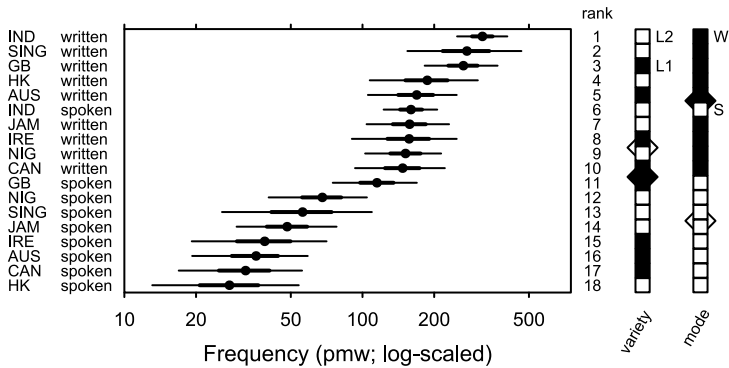
Figure 7.6: Text frequencies: Ranking of specific conditions for *though*; W = written, S =spoken

Finally, Figure 7.7 shows the frequency rankings of all *n* = 18 conditions for *even though*. Like *though*, this conjunction occurs slightly more frequently in L2 varieties (mean rank: 8.8) than in L1 varieties (mean rank: 10.4), and it is also much more frequent in written varieties, with a mean rank of 6.1 as compared to a mean rank of 12.9 in spoken varieties. The lower bound of the range of median values for *even though* is the same as for *though*, but it is more restricted at higher values [28; 112].

Regarding text frequencies, then, the main division in the data runs between spoken and written varieties, while variety status (L1 vs L2) plays no more than a subsidiary role. Differences between L1 and L2 varieties will feature in Chapter 10 as well, but the patterns that emerge there are somewhat difficult to reconcile with what is shown here. The implications will be discussed in §10.3.
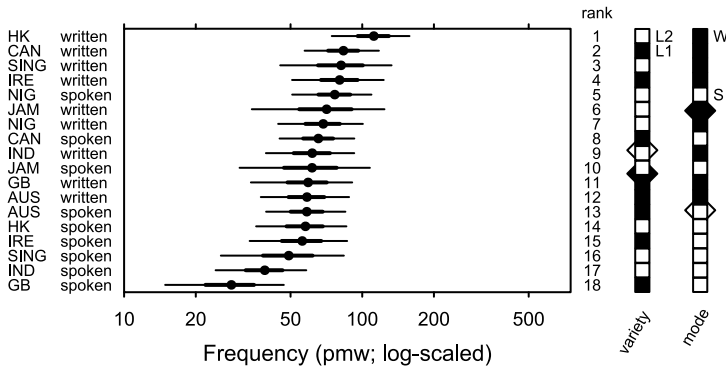
Figure 7.7: Text frequencies: Ranking of specific conditions for *even though*; W = written, S =spoken

7.4 Summary and discussion

The discussion of text frequency patterns in this concluding section will be kept to a minimum and needs to be preceded by a few notes of caution. The purely form-driven investigation of linguistic phenomena based on text frequency has a long tradition in corpus linguistics, possibly because text frequency as such is the most immediately observable and objective aspect in the study of linguistic constructions. The general approach in this study, however, rests on the belief that the construction of a linguistic expression is motivated from the desire to express certain content or certain relations (i.e. the functional side of language), and that it is therefore necessary to treat those functions as predictor variables when analysing or counting forms. This will be the general approach in Chapters 9–11, as explained and shown in schematic form in §4.1.3 above. Against this background assumption, the aim of this chapter has been twofold.

Firstly, it followed up on a notion formulated by Hilpert (2013b: 462; see the very beginning of this chapter): A general investigation of frequencies can be a useful point of departure for the more detailed (or multifactorial) investigation of constructions, since text frequency patterns can be symptoms of underlying cognitive or functional mechanisms and processes. In particular, the frequent exposure to a particular construction type or a particular connective may have consequences for the entrenchment of such linguistic forms. Secondly, findings in this chapter can be juxtaposed with findings in Chapter 10 below, mainly to gauge whether or not the purely form-driven approach can be meaningfully related to analyses that are motivated functionally and consider multiple factors. In many respects, the investigation of text frequency in this chapter will be qualified in the light of the later analyses.

Turning to the results proper, the aggregated frequencies of the three conjunctions for individual varieties fell within a reasonably narrow range, with two notable exceptions: BrE (with a very much higher total rate of use) and NigE (with a very much lower rate). Individual outlier varieties of this kind may arise from the nature and quality of the data (e.g. a lack of corpus comparability due to topic-related sampling error), or there may indeed be a fundamental difference between varieties, in the sense that culture-specific aspects play a role in the use of CCs as discourse-structuring, rhetorical devices. However, explanatory approaches of this kind must at present remain speculative and need to be addressed by independent studies of a different methodological orientation (see comments in Chapter 1).

While the typical frequency ranking in written varieties is $f_A > f_T > f_E$, *even though* is occasionally the second most frequent marker in speech. This is because mode of production has a strong impact on the text frequencies of *although* and *though* (with considerably lower frequencies in speech), while *even though* regularly seems to be immune to this effect. As a result, the role of *even though* in speech is strengthened, relative to the other markers. This has certain implications if we treat mode of production as a basic, binary stylistic variable; the more fine-grained analyses (particularly in Chapter 10) will provide a more detailed picture of this phenomenon.

Finally, let me briefly consider how results relate to the previous research summarised in §5.1.1. The general frequency pattern, with *although* and *though* much more frequent than *even though*, as described, for instance, by Quirk et al. (1985) and confirmed by Altenberg (1986) and Aarts (1988), is also found in the present study. There is, however, not much evidence to support Quirk et al.'s (1985: 1097–1099) treatment of *though* as "more informal" than *although* (see also Biber et al. 1999 and Huddleston & Pullum 2002), except that in many cases *though* responds somewhat less strongly to the spoken/written dimension of variation and thus seems to be less sensitive (or specialised) in this regard. The tendency for *although* to respond most vigorously to a difference in the mode of production confirms a pattern evident in Altenberg's (1986) study based on data from LLC and LOB. Results in the present study also agree with Aarts (1988), who finds that, stylistically, *although* is most sensitive and *even though* is least sensitive in the comparison of the three markers. The emphatic character Quirk et al. (1985) ascribe to *even though* may play a role in this conjunction's higher rates of use in spoken language, if we take a somewhat more generous view on the concept of *emphasis* and extend it to more involved or personal speech styles as found in some kinds of spoken discourse. It must be borne in mind, however, that only *some* spoken genres in ICE can be characterised as involved.

This chapter has shown no more than general patterns, and – with the necessary caveats concerning text frequency analyses in mind – more detailed studies of stylistic variation are called for. There are a few surprising findings in specific varieties of English – see, for instance, patterns in IndE (both spoken and written) and spoken NigE in Figure 7.4. It remains to be seen whether these exceptions are confirmed when more complex, functionally motivated analyses are undertaken in Chapter 10.

8 Frequencies of semantic types

In this chapter, the focus lies on the text frequencies of the four types of intra-constructional semantics that characterise CCs in the present study. As in Chapter 7, the analysis is purely count-based. Whereas the previous chapter inspected forms only (i.e. the three conjunctions), this chapter examines functions only. It is somewhat shorter than the previous one since it contains no analogue of the detailed displays in Figures 7.5–7.7.

One important caveat concerns the limited perspective of the present study with its focus on only three subordinating conjunctions. This chapter therefore cannot truly show whether certain semantic types are generally more frequent in any of the varieties under investigation – it only shows their frequencies in connection with *although*, *though* and *even though*. We could say that these conjunctions constitute a sample of markers, whose representativeness would need to be discussed. Accordingly, it is quite possible that frequency differences as shown in this chapter do not hold true when a wider range – or, ideally, the complete inventory – of concessive markers is taken into account. Therefore, even more so than Chapter 7 above, this chapter generates few insights that can truly stand alone, which is another reason for its relative shortness.

The expectations formulated in §5.3 are that both L2 varieties and the written mode should lean towards higher rates of anticausal CCs (and, perhaps, the related epistemic CCs, too), while dialogic CCs will occur at higher rates in L1 varieties and in speech. Narrow-scope CCs, while they are semantically part of the dialogic class of CCs, cannot easily be captured by the same general hypothesis regarding the effect of mode. They are more integrated as they encode concessive relationships at the phrase level, and they should therefore probably be regarded as cognitively quite complex, making their appearance in writing more likely. The general frequencies of epistemic CCs are expected to fall between those of anticausal and dialogic CCs, due to the alleged intermediate historical status of this type – that is, if we work on the assumption that in the history of English anticausal CCs are primary and develop towards dialogic CCs via epistemic CCs (see discussion in Sweetser 1990). Within the chapter, §8.1 introduces the statistical model for the main frequency analysis in §8.2, and there will be a concluding discussion in §8.3.

8.1 Statistical model

The estimation procedure for rates of occurrence of the four semantic types is similar to the approach in Chapter 7, i.e. a Bayesian negative binomial mixed-effects regression model of identical form is fitted for each variety. Instead of MARKER, however, it is TYPE that is included in the fixed part. This predictor has four levels in this chapter – "anticausal", "epistemic", "dialogic" and "narrow-scope" – while in later analyses (Chapters 9–11) only the two main variants, "anti-causal" and "dialogic", are included. The model is called "Model B", and its syntax is shown in (84); for a detailed specification of variables, see Table 6.3 in §6.3.6. In analogy to Model A above, the two predictors in the fixed part of the model interact, and slopes for TYPE vary randomly across GENRE. Again, the cluster variable TEXT is not included, because the individual text is the smallest unit of observation, and variety does not feature as a cluster variable either, since an independent submodel is run for each variety. As in Model A, the variable LOG_WORDS stands for the logged number of words per text; again, note that this variable is not listed in Table 6.3.

(84)   Model B: Syntax

```
count ~ spoken.ct * type
        + (type | genre)
        + offset(log_words)
```

Appendix B.2 provides more information concerning token numbers, the number of levels of the random factor GENRE, the priors that were specified, and the number of posterior samples. Data, scripts and regression tables can be found in the online materials (cf. §1.4).

8.2 Results

In analogy to Chapter 7, frequencies of the four semantic types in all varieties will first be discussed at a global level before differences between speech and writing are shown. Summing up the text frequencies of all four types would make little sense, as the result would be exactly the same as the summed frequencies of conjunctions (cf. Figure 7.2 in §7.3 above).

Figure 8.1 shows the frequencies of the four semantic types – dialogic, anticausal, epistemic and narrow-scope – in all nine varieties under investigation, based on the respective single-variety components of Model B (see Appendix B.2).

The effects of mode are once again controlled for by estimating average values across written and spoken conditions (cf. §6.3.4 for a discussion). The horizontal arrangement of semantic types in the plots in this chapter only partly follows theoretical considerations as outlined in Chapter 3: Anticausal, epistemic and dialogic CCs are grouped together because they are characterised by the same syntactic flexibility and their subordinate clauses have scope over the entire matrix clause; narrow-scope CCs are set apart, because their syntactic behaviour is more restricted and they have scope over the matrix-clause VP only (cf. §2.2.4). However, within the group of the three central (wide-scope) types, the horizontal arrangement is not based on the putative sequence of their historical development (anticausal > epistemic > dialogic) but follows their typical frequency ranking (f_d > f_a > f_e) to facilitate the comparison of patterns in the plots.
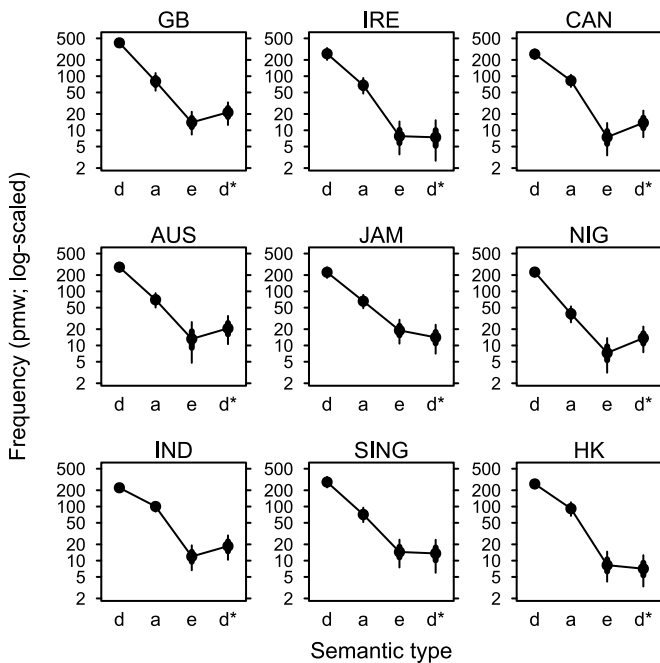
Figure 8.1: Average frequencies of semantic types; d = dialogic, a = anticausal, e = epistemic, d* = narrow-scope dialogic

The literature implicitly suggests that anticausal CCs are somehow primary, or prototypical, possibly because there are relatively straightforward connections between their semantics and those of related kinds of adverbials (e.g. conditional, causal, or consecutive). As shown in Figure 8.1, however, dialogic CCs are the most frequent semantic type in all nine varieties under investigation, with a mean

text frequency of 266 pmw (not shown), followed at a considerable distance by anticausal CCs (M = 72 pmw).[1] For epistemic CCs, M = 11 pmw across all varieties, and for narrow-scope CCs M = 14 pmw. With regard to these latter two types, there is greater variability between varieties: Epistemic CCs are more frequent than narrow-scope CCs in IrE, JamE, SingE and HKE, while the opposite is the case in the remaining five varieties. Since they are semantically related, "regular" dialogic CCs (with scope over the entire matrix clause) and narrow-scope dialogic CCs could alternatively be treated as a single category, with a summed frequency value. However, due to the logarithmic treatment of the frequency scale in Figure 8.1, this would hardly affect the position of dialogic CCs relative to the others.

In the next step, the estimated frequencies of CCs of the four semantic types are compared between spoken and written varieties. Thus, in analogy to Figure 7.4 in the previous chapter, two estimates are shown for each type in each variety in the lower panels of Figure 8.2, complemented by the estimated ratio of the two in the upper panels. In all nine varieties, dialogic, anticausal and narrow-scope CCs are more frequent in writing than in speech, and the writing-to-speech ratio is relatively robust in most cases: All of the 50% intervals and most of the 90% intervals are above the value of Ro = 1. Across varieties, the mean ratios (f_W/f_S) are Ro_d = 2.7, Ro_a = 2.1, Ro_e = 1.6 and Ro_{d^*} = 3.6, respectively. Thus, particularly narrow-scope CCs are much more common in written language. Another way to look at this phenomenon is to compare the rankings of the two less frequent types: In the written mode, narrow-scope CCs are more frequent than epistemic CCs in seven out of nine varieties (exceptions being SingE and HKE); in the spoken mode, this is the case in only three varieties (CanE, AusE and NigE). In contrast to narrow-scope CCs, the frequencies of epistemic CCs differ less markedly between modes of production: There are four varieties (BrE, JamE, NigE and IndE) in which there is virtually no difference.

Although dialogic CCs could be argued to require less planning than anticausal or epistemic CCs, as they are characterised by semantically less tightly connected (or integrated) propositions, there is no evidence that they are more frequent in spoken language. If this were the case, the written/spoken ratios of dialogic CCs should be lower (perhaps even below Ro = 1) when compared to anticausal CCs, for instance. However, text frequencies of dialogic, anticausal and narrow-scope CCs all mirror the general pattern established in the inspection of the frequencies of individual markers.

---

[1] Again, the geometric mean of variety-specific medians was used in this and the following calculations. See Footnote 2 on p. 119.
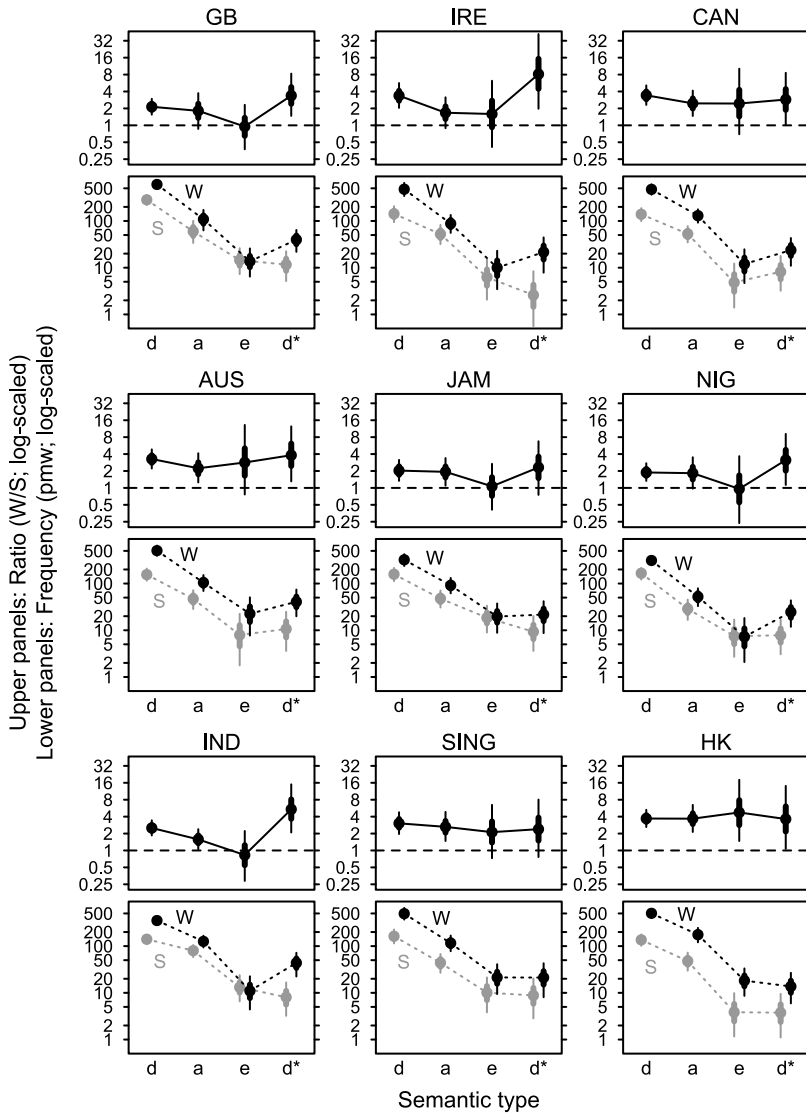
Figure 8.2: Frequencies of semantic types in speech and writing; d = dialogic, a = anticausal, e = epistemic, d* = narrow-scope dialogic, W = written, S = spoken

8.3 Summary and discussion

In the analyses in this chapter, dialogic CCs emerged as by far the most frequent type in all varieties. Examples of CCs typically cited in the literature tend to belong to the anticausal type, which seems to be implicitly treated as the semantic prototype. However, anticausal CCs are only the second most frequent semantic type in the present study. At far lower text frequencies, we find epistemic and narrow-scope dialogic CCs. In some varieties, the former is more frequent than the latter; in others, the opposite pattern obtains. The comparison of spoken and written varieties confirms the frequency patterns discussed in Chapter 7, with generally higher frequencies in writing for all types, perhaps with the exception of epistemic CCs, which are often enough of similar frequencies in both modes. There is no evidence of the dialogic type being more frequent in speech, although it could be argued to be more coordinated in character and thus to require less advance planning. The frequencies of narrow-scope CCs, on the other hand, differ more radically between speech and writing than those of the other types, with considerably higher frequencies in writing. This may be because in most CCs of this type the complement of the conjunction is nonfinite – usually it is not even clausal or interpretable as a verbless clause (see examples in §2.2.4). Moreover, narrow-scope CCs are more locally embedded at the phrase level, and both this and their nonfiniteness make them cognitively complex and thus likely to correlate with written language.

There are no previous findings to which results presented in this chapter could be related, with the exception of Hilpert (2013a) and some of my own earlier research. Although not including *even though*, Hilpert's study suggests that the anticausal type is much more frequent than seems to be the case in the present study. As pointed out earlier, however, Hilpert focuses on a specific construction type (with co-referential subjects in both component clauses), whose semantic versatility is possibly restricted. Schützler's (2017, 2018a) results anticipate the outcome of this chapter, even if, like Hilpert's, these studies take a semasiological approach to markers, investigating the percentages of semantic types found for each of them.

Finally, it has to be mentioned once again that the approach of counting semantic types the way it was done in this chapter has its limitations and is attached to certain caveats. Since the analysis is strictly limited to the three conjunctions *although*, *though* and *even though*, it is not transparent how many of the different semantic types are encoded using other formal means, like prepositional constructions or coordination with *but*, for example. It may well be the case that the total number of CCs (at least of the dialogic, anticausal and epistemic types)

does not differ substantially between speech and writing when inspected from a more holistic, global perspective that takes more concessive markers into account. In other words: Frequencies of semantic types shown in this chapter are to a considerable extent likely to be artefacts of the frequencies of subordinating conjunctions.

9 Clause position

This chapter focuses on the relative frequencies of the two basic configurations of matrix and subordinate clause, "final" and "nonfinal". The nonfinal category comprises sentences with subclauses in medial and initial position (cf. §2.3.1, §5.1.3 and §6.3.6). Unlike the analyses in Chapters 7 & 8, the approach is not based on text frequency counts but on the inspection of variable contexts, i.e. the choice that is made between the two positional variants each time a CC occurs. Results are presented as percentages.

The main expectations concerning the variation of clause positions are that nonfinal clause placement is somewhat more likely in L1 varieties, in written discourse and in connection with anticausal semantics. For the reasoning behind these assumptions, see §5.3, which also highlights the place of positional variation in the choice model of constructional variation for CCs (see Figure 5.11). Similar in general structure to the foregoing ones, this chapter will first present a minimal discussion of how the data were approached in the statistical model (§9.1), followed by the results in §9.2. Finally, §9.3 focuses on a concluding summary and discussion of the quantitative findings.

9.1 Statistical model

Clause positions are classified as either final or nonfinal, and the outcome is thus a binary variable called FINAL, with the reference level "nonfinal" comprising initial and medial positions (cf. §2.3.1; for a detailed definition of variables, see Table 6.3 in §6.3.6).[1] Accordingly, a binary logistic regression model was used to estimate the outcome. In line with the sequence of chapters, this is called "Model C"; its structure is shown in (85). Again, one model was specified for each variety, resulting in a total of $n = 9$ models. The interaction term for the predictors SPOKEN.CT and ANTI.CT was included; further, slopes of ANTI.CT were specified as varying randomly across the two cluster variables GENRE and TEXT.

---

[1]Note that the (uncentred) outcome variable FINAL is different from the (centred) predictor variable FINAL.CT (cf. Table 6.3).

Note that, as in all models in this study, there is no cluster variable for variety, since the nine varieties were each assigned a separate model. The predictor LENGTH.CT is used as a control variable. It measures the logged length (in words) of subordinate clauses, centred on the mean logged length of all occurrences, across all varieties (see §6.3.6). This predictor will play no prominent role in the discussion of results.

(85) Model C: Syntax

```
final ~ spoken.ct * anti.ct + length.ct
       + (anti.ct | genre)
       + (anti.ct | text)
```

More information can be found in Appendix B.3, e.g. concerning token numbers, the number of levels of both random factors (GENRE and TEXT), the priors (which were the same for all nine models) and the number of posterior samples. Data, scripts and detailed model summaries can be retrieved from the online repositories as outlined in §1.4.

9.2 Results

The three parts of this section present the results from two different perspectives, first averaging across varieties (§9.2.1), then inspecting the relationship between intra-constructional semantics on the position of the subordinate clause (§9.2.2). Both sections also take the spoken-written dimension into account.

Before embarking on the main analyses, a few words about the effect of clause length (operationalised as the predictor LENGTH.CT) are in place. This predictor was used purely as a control variable: The effect of length on the placement of syntactic elements (or on syntactic variation more generally) is well documented, and while the present study takes no theoretical interest in it, ignoring the effect would have been problematic. While the predictor LENGTH.CT is therefore included in all the models, it is held constant at its mean value. Values of this variable are positively correlated with the outcome FINAL, as seen in Figure 9.1, which plots the coefficients (logits) for all nine varieties, ordered by magnitude and including 50% and 90% uncertainty intervals.[2] While the interpretability of these values is limited, they show that the effect is always positive, and sometimes substantially so.

---

[2]The coefficient signifies the change in the log odds of final position corresponding to a unit change of LENGTH.CT, which in turn corresponds to a change in actual length (measured in words) by factor $e \approx 2.72$.
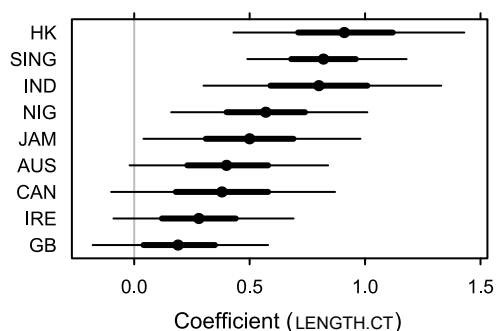
Figure 9.1: The coefficient LENGTH.CT in nine varieties

Compared to a preliminary model run without the predictor LENGTH.CT, the coefficients for the intercept, SPOKEN.CT and ANTI.CT in the fixed part changed only very slightly, in the vast majority of cases by no more than ±0.04 on the log odds scale. Random coefficients tended to increase in the model that included LENGTH.CT as a predictor. On closer inspection, it emerged that this predictor was neither correlated with any of the others, nor with the individual varieties. Using a reduced model without LENGTH.CT would therefore have yielded only marginally different results and would certainly not have affected the general conclusions.

9.2.1 Average percentages

The first approach to the positioning of subordinate concessive clauses relative to their associated matrix clauses was to establish average values for individual varieties. This perspective makes the perhaps questionable assumption of neutral values for mode of production (intermediate between spoken and written) and semantics (intermediate between anticausal and dialogic), but it is nevertheless informative and provides an easy-to-grasp point of departure for subsequent analyses (see discussion in §6.3.4). Figure 9.2 shows percentages of subordinate clauses in sentence-final position for all nine varieties under investigation, arranged in ascending order. L1 varieties are shown in black, L2 varieties in grey; the central dashed line represents the 50% mark, which can be used to assess whether or not the data encourage the conclusion that either final or nonfinal clause placement can actually be considered the majority variant.

Values range from 39.5% of clauses in final position in JamE to 55.9% in CanE. The three varieties on the left (JamE, IndE and HKE) show a very clear preference of nonfinal placement; the other varieties are considerably closer to the 50% value.
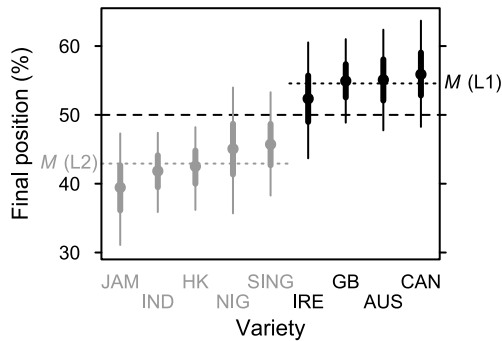
Figure 9.2: Average percentages of subordinate clauses in final position

Further, all L2 varieties prefer subordinate clauses in nonfinal position, while L1 varieties prefer final placement. In Figure 9.2, this difference is emphasized not only by the left-to-right orientation of the two groups, but also by the indication of means for L1 and L2 varieties, based on the respective variety-specific averages: For L2 varieties, the average percentage of sentence-final subordinate clauses is 42.9%, while for L1 varieties it is 54.6%, a difference of 11.7 in absolute percentage points.

As mentioned above, the effects of the mode of language production (spoken vs written) as well as the intra-constructional semantics of concessives are neutralised – or controlled for – in Figure 9.2. A leading question in the subsequent analyses will be whether or not the general difference between L1 and L2 varieties detected here unfolds into further differences concerning specific effects. In other words: Are the L1 and L2 varieties investigated here characterised by different general preferences concerning clause placement but otherwise affected similarly by differences in the mode of production or the semantic type of a concessive, or do they also respond differently to those factors?

The first step towards a more nuanced assessment of differences in the preferred clause placement patterns is taken in Figure 9.3, which again orders varieties according to their average percentage of sentence-final subordinate clauses. However, each variety category in the lower panel is now subdivided into "spoken" and "written", shown in grey and black, respectively. Once again, a line is drawn at the value of 50% since values that depart more markedly from this reference point indicate that there is an actual preference in a given (spoken or written) variety. The upper panel shows estimates of absolute percentage-point differences between speech and writing in each variety.

Percentages of subordinate clauses in final position tend to be higher in spoken language in most varieties. The contrast – written minus spoken – takes a (rather small) positive value in only two varieties, SingE and BrE. Differences between
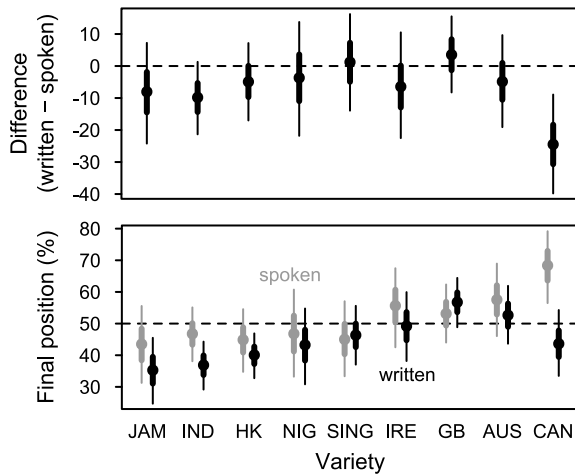
Figure 9.3: Sentence-final placement of subordinate clauses by mode of production

modes of production generally tend to be moderate and come with a high degree of uncertainty – observe, for instance, that the 90% uncertainty intervals include the critical value of zero in all varieties except CanE.[3] If we inspect the average differences between speech and writing across subsets of varieties, we see that L1 and L2 varieties behave similarly in this regard, even if their overall percentages differ substantially (as shown in Figure 9.2): Compared to writing, the mean share of subordinate clauses in final position in spoken L2 varieties is on average higher by 5.0 percentage points; in L1 varieties, this average difference is 8.1 absolute percentage points, as summarised in Table 9.1.[4]

It follows that Table 9.1 also shows that the general difference between L2 and L1 varieties – with the former characterised by fewer subordinate clauses in sentence-final position – persists in both modes of production. From the bird's-eye perspective discussed above, this difference was 11.7 absolute percentage points. In speech and writing, it is relatively similar, at 13.3 and 10.2 percentage points, respectively.

---

[3]Readers are invited to inspect the regression tables that are published online (cf. §1.4), which show, for instance, that CanE not only has a high INTERCEPT for final position – reflected here in its position at the very right of the figure – but also an exceptionally high coefficient for SPOKEN.CT. Consulting the supplementary materials in this way is generally recommended to readers who do not wish to rely exclusively on the visualisations.

[4]Note that only the mean values themselves are directly derived from the model-based estimates. Differences in the table are then calculated on this basis to present a consistent picture: If, for the rightmost column, variety-specific mean differences between speech and writing were estimated and then averaged for each group, slight and uncritical (but potentially confusing) discrepancies might arise. This also applies to the corresponding tables in Chapters 10 & 11.

Table 9.1: Sentence-final placement of subordinate clauses in speech and writing by variety type (mean %)

	written	spoken	written − spoken
L2	40.4	45.4	−5.0
L1	50.6	58.7	−8.1
L2 − L1	−10.2	−13.3	

9.2.2 Semantics

In an approach strictly analogous to the one taken in the previous section, the lower part of Figure 9.4 isolates the effect of the intra-constructional semantics of a CC – with dialogic and anticausal types shown in grey and black, respectively – on the placement of subordinate clauses. In the upper panel, the differences between the two conditions are shown. The same ranking of varieties as in §9.2.1 is applied. Once again, lines of reference are drawn at 50% in the lower panel and at zero in the upper panel.
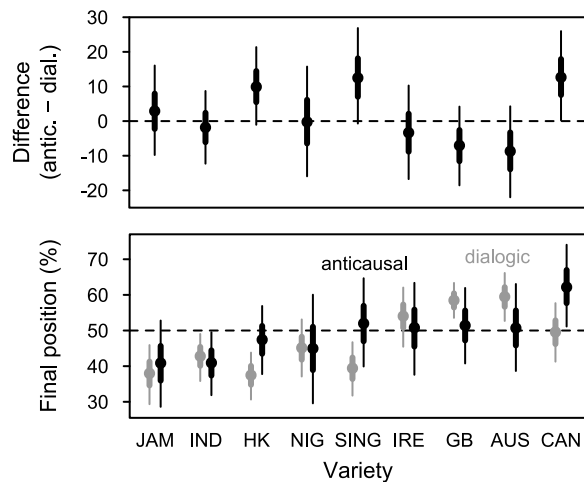
Figure 9.4: Sentence-final placement of subordinate clauses by semantic type

The relationship between clause positions and semantic types does not appear to be systematic. In four varieties (JamE, HKE, SingE and CanE), there is a tendency for anticausal semantics to be associated with higher percentages of concessive clauses in final position, although this effect is rather small in JamE;

in four varieties (IndE, IrE, BrE and AusE), the inverse pattern obtains, with relatively weak or uncertain patterns in IndE and IrE; and in NigE there is virtually no difference, the expected percentage of subclauses in final position being lower by a mere 0.23 percentage points for anticausal CCs. Overall, there is no clear, general tendency with a few exceptions (as in the analysis of speech vs writing above) but a mix of indifferent or even conflicting patterns.

In the following paragraphs, the relationship between semantics and clause positions will be explored in more detail by including the spoken-written dimension as a superordinate level of variation – in other words, the interaction of the predictors SPOKEN.CT and ANTI.CT is taken into account. Results are shown in Figure 9.5, which requires a few words of introduction. For each of the nine varieties, there is one component plot with two panels (lower and upper). Varieties are no longer ordered by median percentages but according to the sequence introduced in §4.3 and §6.1 (cf. Table 4.1 and Figure 6.1). The lower panel of each plot shows the estimated percentages of subordinate clauses in final position in the four possible conditions (2 modes × 2 semantics), while the upper panels show the difference between anticausal and dialogic semantics in speech and writing.

The very first component plot (representing BrE) shows the level of detail that may be revealed by including the interaction between mode of production and semantics: In speech, anticausal concessives are considerably less likely to be constructed with a sentence-final subordinate clause (−21.5 absolute percentage points), while in writing there is a less pronounced tendency in the opposite direction (+7.4 percentage points). There are multiple strategies for reading Figure 9.5. Focusing on the plots of differences (i.e. the upper panels in each subplot), point estimates below the dashed line signal that, compared to anticausal CCs, dialogic CCs have a higher percentage of subordinate clauses in final position. This is the expected outcome (cf. §5.3) on the assumption that clause arrangements should be iconic of the underlying IF→THEN relation in anticausal CCs, but should follow patterns that are easier to parse in dialogic CCs. On the other hand, if the point estimate is above zero, a higher percentage of subordinate clauses in final position is associated with anticausal CCs, which is contra expectations. Secondly, if the connecting line in a plot of difference (upper panels) has a relatively flat slope, or is even parallel to the *x*-axis, the effect of intra-constructional semantics on clause positions is similar in speech and writing. If there is a marked difference, i.e. if the connecting line slopes steeply, the semantic effect differs substantially between modes of production.

The plot confirms what the previous section has shown, albeit in more detail: The response of clause arrangement to the semantic predictor is relatively erratic and unsystematic. Additionally, there is often a difference in the semantic effect
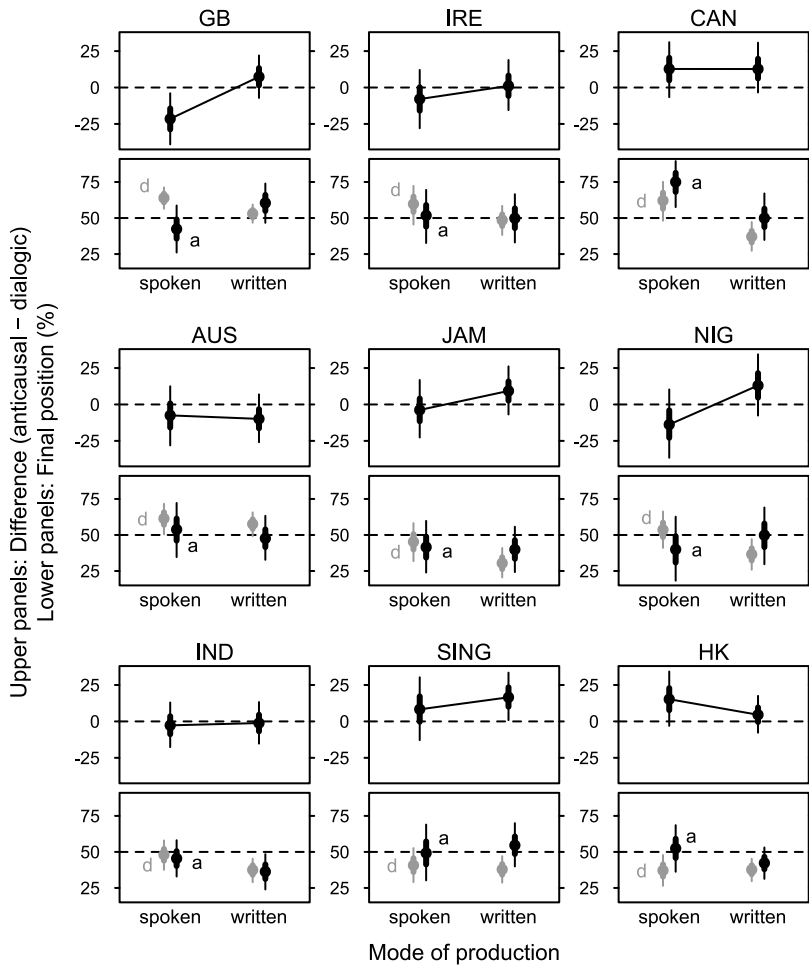
Figure 9.5: Clause position: The interaction of mode and semantics; a = anticausal, d = dialogic

between the two modes of production that is equally surprising and difficult to explain. The unexpected result found in BrE was already discussed above, with the expected pattern in speech but no effect in writing; IrE has an unexpected effect in speech and virtually no effect in writing; CanE is characterised by tendencies (in both modes of production) whose directions run counter to hypotheses; AusE seems to conform to the hypothesised patterns and is stable across both modes of production, even if the effects are not particularly large; in JamE, there is virtually no effect in speech but an unexpected tendency in writing; in NigE, there is a tendency in the expected direction in spoken language, but its complete reversal in writing; IndE shows virtually no semantic effect, irrespective of the mode of production; both SingE and HKE have effects in the expected direction, but their strength varies between speech and writing in an inconsistent way. In short: There is no evidence to support the idea of an iconic arrangement of matrix and subordinate clause relative to each other, and the balance of evidence and counter-evidence discourages further interpretation. Clause positions thus cannot be explained using the model that was proposed for this chapter. I will return to this finding in the concluding section and offer a few suggestions for future approaches to the issue.

Figure 9.6 is similar in design to Figure 7.5 and provides a final visual summary for this chapter. In its main panel, it does not contain information that goes beyond what was shown in Figure 9.5 above but arranges values in a different, cross-varietal manner (cf. Figures 7.5–7.7). Ranked percentages of subordinate clauses in final position are plotted for all $n = 36$ possible conditions (9 varieties × 2 modes of production × 2 semantics). The black, white and grey boxes to the right of the figure highlight structure in the data. In the first three columns, black squares denote L1 varieties, written language and anticausal semantics, respectively; conversely, white squares denote L2 varieties, spoken language and dialogic semantics. In the fourth column, three categories are established, defined by the interaction of mode and semantics (cf. Figure 9.5). Additionally, the mean ranks of groups of conditions – "black" vs "white" (vs "grey") – are indicated in each column by triangles that jut out to the left and right. Darker colours in the right-hand part of the figure correspond to conditions that should favour the nonfinal placement of subordinate clauses (L1 varieties, writing and anticausal semantics), according to the hypotheses. A concentration of darker shades nearer the bottom of each column (and corresponding mean ranks) would therefore signal agreement between results and expectations.

The reversal of the expected varieties-based pattern is clearly visible in the clustering of black squares towards the top of the first column in the right-hand part of Figure 9.6 (mean rank: 11.6) and the higher concentration of L2 varieties
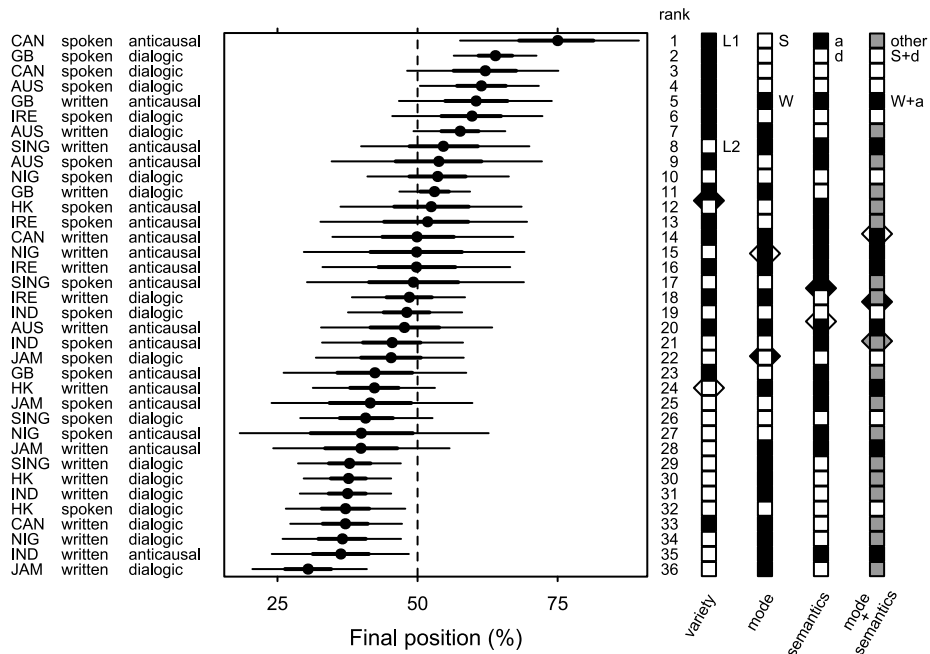
Figure 9.6: Clause position: Ranking of specific conditions; W = written, S = spoken, a = anticausal, d = dialogic

towards the bottom (mean rank: 24.1). The general effect of mode of production is as expected: White squares representing speech in column two are on average nearer the top (mean rank: 15.1) than black squares representing writing (mean rank: 21.9), but there is a high degree of overlap between the two sets of specific conditions. Concerning the difference between anticausal and dialogic CCs, Figure 9.6 – like Figure 9.4 above – shows that it has no systematic impact on the positioning of clauses, as the mean ranks for both types are very close to each other (17.4 and 19.6, respectively). Finally, the pattern seen in the column at the very right of Figure 9.6 highlights once again that hypotheses concerning the sequencing of clauses are not, or only very partially, supported. While dialogic CCs in speech are indeed most likely to be associated with subordinate clauses in final position (mean rank: 13.8), the intermediate combinations of factors (spoken anticausal and written dialogic) have a mean rank of 20.9, which is lower than for the hypothetically most strongly disfavouring combination, written anticausal (mean rank: 18.3).

There are thus many patterns that are unsystematic (or noisy) or even run counter to the hypotheses that were set up to account for alternating clause posi-

tions. Except for the difference between speech and writing, results match poorly with theory. At present, it can therefore only be concluded that (i) the theoretical assumptions for this part of the investigation may not be adequate, particularly concerning the iconicity principle, or that (ii) there are other (and potentially stronger) factors at work that were not operationalised for this study. One such candidate factor will be discussed in the following final section of this chapter.

9.3 Summary and discussion

The investigation of factors that correlate with the final or nonfinal placement of subordinate clauses in CCs yielded mainly three results: (i) Sentence-final position is more common in L1 varieties than in L2 varieties, (ii) spoken language correlates with subordinate clauses in final position, and (iii) there is no systematic general link between the semantic relation that holds within a CC and the arrangement of its component clauses relative to each other. Only the second finding is in support of the corresponding hypothesis formulated in §5.3. In the following paragraphs, I will briefly discuss these rather ambivalent results and speculate as to the reasons for their lack of coherence. Ultimately, I will argue that clause position as an outcome variable is inherently problematic, at least in an analytic design that focuses on hermetic constructions and ignores the wider discourse context.

Concerning the link between types of varieties (L1 vs L2) and clause placement, the initial hypothesis was that it would be L2 varieties that favour subclauses in final position. As discussed in §2.3.1, theories connected to production and parsing suggest that final placement is cognitively the optimal configuration. In L2 varieties, the role of English will on average be somewhat less secure, compared to L1 varieties, and its share in everyday language use will be smaller. Under such conditions, it was argued, the selection of cognitively less complex (or more "natural") patterns would be more likely. However, the opposite seems to be the case in the data at hand, as it is the L1 varieties that are characterised by more subordinate clauses in final position. Cognitive mechanisms in language production and processing cannot provide an explanation. It is tempting to resort to a post-hoc inversion of the hypothesis. For instance, one could build a two-stage argument based on the assumption that concessive subordinate clauses in initial position are more frequent in L2 varieties due to the way in which language is acquired: (i) Many standard grammars of the language tend to focus on what may be seen as prototypical CCs, i.e. anticausal semantics with a preposed subordinate clause, and (ii) the acquisition of L2 Englishes may be viewed as more "scholastic", i.e.

happening to a much greater extent in formal school settings, which depend on input from such grammar books and derived materials. L1 Englishes, on the other hand, could then be viewed as more emancipated from what is codified in grammars, which would allow it to follow the more "natural" tendencies predicted from a production or parsing perspective.[5] However, exploring this alternative set of hypotheses would require a new, independent research effort, probably incorporating theories of language acquisition, an inspection of teaching materials and practices, and possibly experimental techniques.

The second result concerns the relationship between modes of production and the placement of clauses. There is a fairly consistent tendency in the data for spoken language to favour subordinate clauses in final position, even if the effect is not particularly strong (with the exception of CanE). This finding is in line with the hypothesis outlined in §5.3: From a production-and-processing perspective, final placement was considered to be cognitively less demanding than nonfinal placement and was therefore expected to be favoured even more strongly when the linguistic signal is purely acoustic and thus transient. The finding that speech tends to favour final clause placement also agrees with Altenberg's (1986) results.

Thirdly and finally, the association between semantics and clause position in this study does not follow a clear and interpretable pattern. There is no support for the hypothesis that the arrangement of clauses in anticausal CCs should be iconic of the intra-constructional semantic relation between propositions (again, see §2.3.1 and §5.3). Individual patterns that confirm the hypothesis co-occur with patterns that run counter to it, so that the overall picture is very difficult to interpret.

In view of the results presented in this chapter, there remains a feeling of unease with the treatment of clause position as an outcome variable. One obvious general conclusion could be that the factors operationalised for the analysis are not the centrally important ones. In other words, they may at the very least be obfuscated, if not outright overridden, by other determinants not even considered here. One such factor with a potentially strong effect on the sequencing of matrix and subordinate clauses is information-structural in nature. This means that the particular arrangement of clauses in a CC depends on which proposition SP/W wishes to place in focus position in order to give the sentence as a whole a specific theme-rheme (or topic-comment) structure. This decision, it can be assumed, will be partly subjective but probably to a larger part determined by the wider

---

[5]My observation concerning the predominance in grammars of anticausal CCs with preposed subordinate clauses is largely impressionistic and has not been tested systematically. This further undermines the alternative hypothesis, in addition to the fact that it is post hoc.

discourse context and the pragmatic function of the entire CC within it. In other words, the conditioning factor may in this case be external to the construction itself, at least as defined in the present study. Conceivably, such information-structural mechanisms may be stronger than (and, of course, independent of) factors related to production and processing, or the iconic relationship between syntactic and semantic structures. Thus, in order to understand better why a particular clause arrangement is selected, we may need to look beyond the CC and inspect its discourse function relative to what follows and goes before. If it is a rather taxing exercise to classify CCs as belonging to one of the categories established for this study, operationalising the wider discourse context would be even more involved. As discussed in §1.1, a discourse-analytic approach was explicitly not taken in this study, and the decision to conduct analyses entirely at the level of the CC itself was made to enable the quantitative approach.

Thus, as far as positional variation is concerned, the success of the analysis is limited. Disappointing though this may be, the findings from this chapter are in fact valuable pointers for future research on concessives and perhaps other types of adverbials. Crucially, while the analyses in this chapter provide only limited insights into clause position as an outcome variable, this does not automatically disqualify it as a predictor variable for subsequent analyses. The choice model introduced in §4.1.3 (see Figure 4.2 there) is rather tolerant regarding explanatory gaps: We can accept that our understanding of why SP/W selects a certain clause arrangement remains limited, perhaps because we have given insufficient consideration to additional predictor variables, or because variation is to a large extent unsystematic. In spite of this, we can still use clause position (along with mode of production and semantics) as a predictor in subsequent stages of the study.

10 Choice of conjunction

In this chapter, the focus lies on factors that affect the choice between the three conjunctions *although*, *though* and *even though*. This perspective could be argued to take centre stage in the study as a whole, since specific conjunctions constitute concrete morphological forms and are therefore perhaps more immediately noticeable (or salient) in a CC than, for example, semantic structures or clause positions – they are, after all, the connecting devices upon which a CC hinges. Predictors used at this stage of the analysis follow from the choice model presented in Figure 4.2 (see §4.1.3) and include the mode of production, the semantic type of a CC and the internal arrangement of clauses. Section 10.1 defines the statistical model used in this chapter, followed by a presentation of results in §10.2. The summary in §10.3 reflects upon these results against the background of the expectations formulated in §5.3.

10.1 Statistical model

The outcome variable MARKER takes three values: *although* (the reference category), *though* and *even though*. Thus, "Model D" – shown in (86) – is a multinomial mixed-effects model (cf. §6.3.3.3). The two fixed-effects terms ANTI.CT and FINAL.CT interact with SPOKEN.CT but not with each other. That is, the effects of semantics and clause position on the selection of the marker may differ between speech and writing but are treated as independent of each other. Both ANTI.CT and FINAL.CT vary randomly across the two grouping factors GENRE and TEXT, which are the same as in Model C (see §9.1 above). As in all other chapters, a separate model of the same syntax was fitted for each variety.

(86)   Model D: Syntax

```
marker ~ spoken.ct * (anti.ct + final.ct)
        + (anti.ct + final.ct | genre)
        + (anti.ct + final.ct | text)
```

Appendix B.4 contains information regarding token numbers, the number of levels of both random variables (GENRE and TEXT), the priors (constant across all nine models) as well as the number of posterior samples. For data, scripts and model summaries (i.e. tables with regression coefficients), see the online repositories (cf. §1.4).

10.2 Results

Results are presented in four sections. The first three of these (§10.2.1–10.2.3) take the following perspectives: (i) a global one in which the effects of both semantics and clause position are controlled for (cf. §6.3.4), (ii) one in which the focus lies on the effects of the two semantic types (controlling for positional effects), and (iii) one in which the focus lies on positional effects (controlling for semantic effects). In each case, a hypothetical, average scenario (poised between speech and writing) is given first, followed by one that takes the effects of mode of production into account. Finally, §10.2.4 documents the full, most detailed range of results by showing for each conjunction the $n = 72$ specific conditions that affect its probability of occurrence.

10.2.1 Average percentages

As in the foregoing chapters, the first perspective on the outcome – in this case the estimated percentages of the three concessive conjunctions – is based on global averages in the nine varieties under investigation. As pointed out earlier, the approach is hypothetical in suggesting that semantics can be indeterminate between anticausal and dialogic, and that clause positions can be indeterminate between final and nonfinal. However, showing all factors in combination and thus applying no generalisation and simplification would seriously hamper the understanding of individual effects.

Figure 10.1 displays average percentages of the three conjunctions *although*, *though* and *even though* in the nine varieties from a global perspective. The arrangement of varieties follows the sequence in Figure 6.1 (and Table 4.1); the horizontal arrangement of conjunctions in each panel is in accordance with the order in which the markers were introduced in the theoretical part. Connecting lines are added to facilitate the direct comparison of patterns.

Typically, varieties are characterised by a "hockey-stick" pattern: *Although* is expected to be most commonly selected ($M_A = 47.4\%$), while values for *though* and *even though* are much lower and roughly on the same level ($M_T = 25.0\%$ and
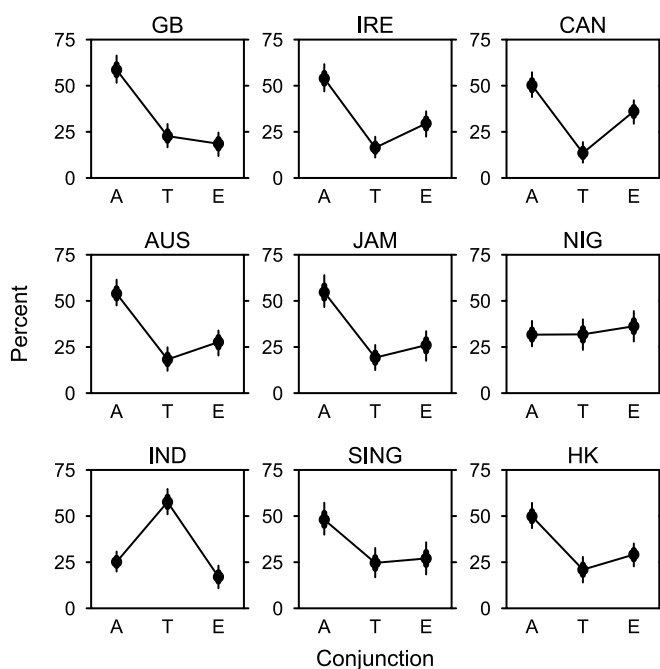
Figure 10.1: Average percentages of conjunctions; A = *although*, T = *though*, E = *even though*

M_E = 27.5%, respectively). There is considerable variation between varieties, however: For *although*, extreme values are found at 58.6% in BrE and 25.2% in IndE; for *though*, the range of values is between 57.6% in IndE and 13.5% in CanE; and for *even though*, extremes are at 36.3% in NigE and 17.1% in IndE. The most striking patterns are found in NigE, which does not seem to give precedence to any of the three markers, as well as in IndE with its remarkably high value for *though* (and, accordingly, a low value for *although*). Both patterns can also be seen in Figure 7.3, although this is based entirely on text frequency and does not control for semantics and clause position. Among the L1 varieties, CanE stands out slightly in using a relatively high percentage of *even though*. Compared to the purely count-based analysis in Chapter 7, the conjunction *even though* has considerably greater weight in Figure 10.1, appearing in second place (after *although*) in six out of the nine varieties and even being the preferred marker in NigE. As will be shown below, there are two main reasons for this: (i) There is a strong positive correlation between anticausal semantics and the use of *even though*, and (ii) anticausal semantics are on the whole considerably less common than dialogic semantics. Since the initial analysis in this chapter assumes a

balance between the two semantic types, the estimated percentages will be positively biased for *even though* and negatively biased for *although*, compared to actual rates of occurrence. Unfolding this general picture into a perspective that does consider semantics as a factor is therefore all the more important, as will be shown in §10.2.2.

Some evidence of a general difference between L1 and L2 varieties is produced by the inspection of the mean percentages of the three conjunctions for those two subgroups. On average, *although* occurs 54.2% of the time in L1 varieties, but only 41.9% of the time in L2 varieties. The respective values for *though* are 17.7% (L1) and 30.9% (L2), while for *even though*, mean percentages in both subgroups are very similar (L1: 28.0%; L2: 27.1%). It is tempting to make conjectures concerning possible explanations of this pattern (e.g. from grammaticalisation theory), and some such notions will be touched upon in the concluding part of this chapter (§10.3), but we need to bear in mind that much of the difference between *although* and *though* is due to the rather idiosyncratic and extreme behaviour of a single variety, IndE. Thus, it seems risky to make even tentative generalisations.

The global perspective in Figure 10.1 becomes more nuanced in Figure 10.2, which is arranged along the same general lines but compares separate percentages of markers for spoken and written genres. Accordingly, there are two sets of values in the lower panel of each variety-specific subplot, rendered in grey and black and internally connected with lines to facilitate the recognition of patterns. In the panels above the percentage plots, differences between writing and speech (in absolute percentage points) are plotted for each conjunction. The relevant reference value of zero (that is, "no difference") is highlighted by a dashed line. In the discussion of tendencies for the individual conjunctions, varieties are ordered according to effect sizes, but note that large effects may also come with high degrees of uncertainty, as indicated in the text and visible in Figure 10.2.

The conjunction *although* tends to be selected more often in written language. There are patterns of this kind in CanE (D = 16.8 [5.3; 29.1]), AusE (D = 12.3 [1.0; 24.1]), IndE (D = 9.1 [−0.4; 19.0]), NigE (D = 7.2 [−5.9; 19.6]), IrE (D = 6.2 [−6.2; 19.0]), JamE (D = 3.1 [−11.3; 17.9]), HKE (D = 2.2 [−9.8; 13.7]) and SingE (D = 1.1 [−14.7; 16.1]). Only in BrE is the tendency reversed, with an extremely small increase in the percentage of *although* in speech (D = −0.8 [−12.8; 11.4]). That is, in eight out of the nine varieties under investigation there is a tendency for *although* to be more frequent in writing. However, based on the uncertainty intervals shown in Figure 10.2 we can speak of a more robust effect in only two of them, CanE and AusE, perhaps with the addition of IndE.

The conjunction *though* is also generally more likely to be selected in writing, compared to speech, namely in HKE (D = 11.3 [0.4; 21.7]), JamE (D = 9.4 [−0.3; 19.6]), CanE (D = 7.6 [0; 15.4]), SingE (D = 6.9 [−6.8; 20.0]), AusE (D = 6.1 [−3.4;
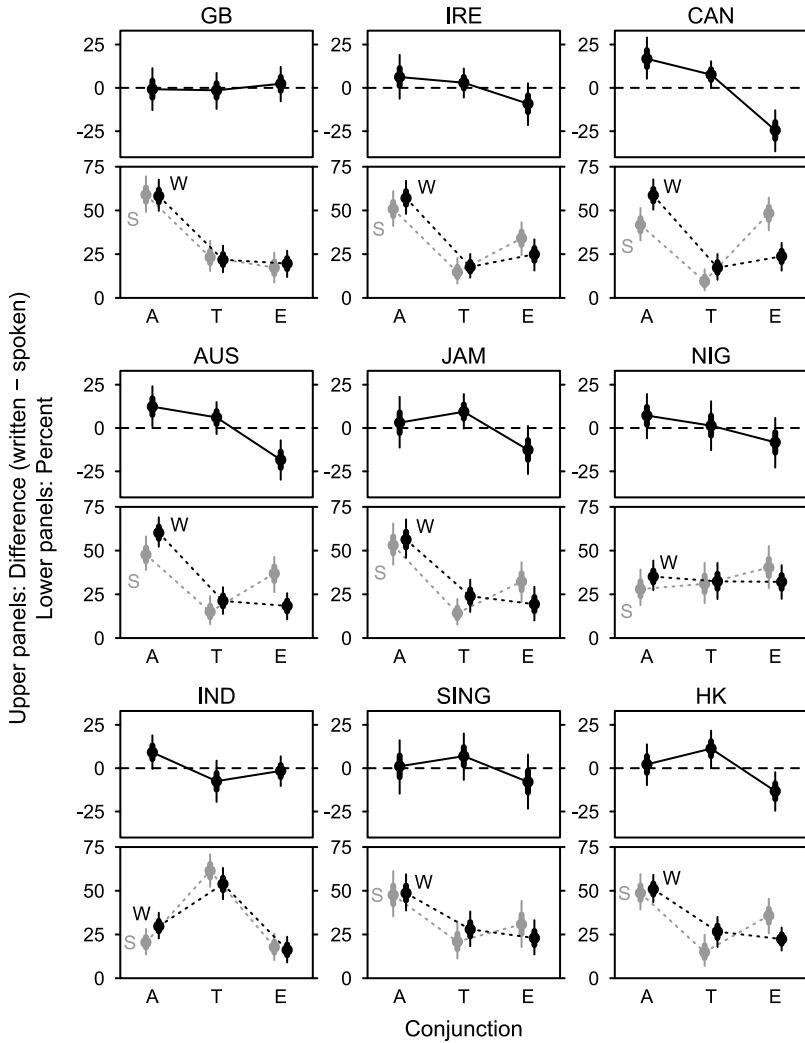
Figure 10.2: Average percentages of conjunctions in speech and writing; A = *although*, T = *though*, E = *even though*, W = written, S = spoken

15.0]), IrE (D = 3.0 [−5.6; 11.1]) and NigE (D = 1.4 [−12.8; 15.4]). A slight reversal is once again found in BrE, i.e. in this variety *though* tends to be selected more often in spoken language (D = −1.4 [−12.1; 8.6]). A more substantial preference of this conjunction in speech is found in IndE (D = −7.5 [−19.4; 4.3]). Not unlike *although*, percentages of the conjunction *though* are higher in writing in seven out of the nine varieties. However, the effect is substantially different from zero in only two of them, HKE and CanE.

In marked contrast to the other two conjunctions, *even though* is more common in speech. This is the case in CanE (D = −24.4 [−36.6; −12.9]), AusE (D = −18.3 [−29.8; −7.1]), JamE (D = −12.6 [−26.5; 1.2]), HKE (D = −12.2 [−24.5; −2.3]), IrE (D = −9.1 [−21.5; 2.7]), NigE (D = −8.3 [−22.8; 5.8]), SingE (D = −7.9 [−23.4; 7.8]) and IndE (D = −1.5 [−10.3; 6.9]). It is only in BrE that we find an effect in the opposite direction (D = 2.4 [−7.8; 12.1]). Thus, from a general, cross-varietal perspective, eight varieties conform to the mainstream tendency for *even though* to be more frequent in speech relative to the other two conjunctions.

Finally, the general differences between the three conjunctions are captured if we average the written-spoken differences across all nine varieties: For *although*, the mean difference in absolute percentage points between writing and speech is +6.4; for *though*, the mean difference is +4.1; and for *even though* it is −10.3. While exceptions do of course exist, we can tentatively conclude that *although* and *even though* are most sensitive to differences in mode of production, and by extension perhaps also to stylistic differences more generally (cf. §4.2).

Table 10.1 summarises the global differences between L1 and L2 varieties concerning the effect of mode of production on the selection of subordinators. It is organised so as to show, for each subset of varieties, the mean percentage of each conjunction in written and in spoken discourse, as well as the difference between those means (written minus spoken) in absolute percentage points (see comment in Footnote 4 on p. 141 regarding this as well as Tables 10.2–10.5 below).

With regard to *although* and *even though*, L1 varieties are on average characterised by a larger percentage-point difference between writing and speech, in a positive direction for *although* and in a negative direction for *even though*. For *though*, the difference between variety types is much smaller. Due to the small number of varieties included in this study, we cannot draw very strong conclusions based on this finding. It is, however, in agreement with the idea that, unlike L2 varieties, L1 varieties have progressed to the differentiation stage in Schneider's (2003) Dynamic Model (see §4.3.2): At a very general level, the greater similarity of percentage patterns in written and spoken L2 varieties may suggest that these two (admittedly very broad) stylistic categories are formally not differentiated to the same extent as in L1 varieties. We will return to this idea in the conclusion to this chapter.

Table 10.1: Mean percentages of conjunctions in written and spoken L1 and L2 varieties

		although	*though*	*even though*
L1	written	58.5	19.5	21.7
	spoken	49.8	15.6	34.2
	written − spoken	8.7	3.9	−12.5
L2	written	44.1	33.0	22.6
	spoken	39.6	28.5	31.4
	written − spoken	4.5	4.5	−8.8

10.2.2 Semantics

In an approach analogous to the one taken in Figure 10.2 in the previous section, the lower panels of Figure 10.3 isolate the correlation of the two semantic categories – dialogic and anticausal (shown in grey and black, respectively) – with the selection of conjunctions. In the upper panels, differences (in absolute percentage points) between the two conditions are shown, subtracting estimated percentages in dialogic CCs from estimated percentages in anticausal CCs. Once again, dashed lines of reference are drawn at the value of zero (denoting "no difference") in the upper panels.

In dialogic CCs, the typical ranking of conjunctions is *although > though > even though*. This pattern is in line with the general frequency pattern described in Chapter 7, and it is explicable from the fact that dialogic CCs are the dominant type – the pattern typical of dialogic semantics will thus have a disproportionately high influence on general text frequencies. Once again, however, IndE with its exceptionally high relative frequency of *though* is a striking exception. Further, *though* and *even though* are roughly on a par in CanE (even with a slightly higher frequency of the latter), and in NigE the percentages of *although* and *though* are almost the same.

Within the category of anticausal CCs, estimated percentages of *even though* in Figure 10.3 are astonishingly high when compared against the initial impressions gained from Chapter 7: In five varieties, this conjunction is the most frequent one of the three, namely in CanE (54.5% [43.9; 64.9]), NigE (47.0% [32.9; 61.4]), AusE (42.1% [30.5; 52.8]), HKE (41.0% [31.4; 50.5]) and SingE (39.6% [26.5; 52.7]); in another three varieties, it ranks second only to *although*, namely in IrE (45.0% [33.2; 56.1]), JamE (37.2% [25.1; 48.6]) and BrE (32.7% [20.8; 43.3]).
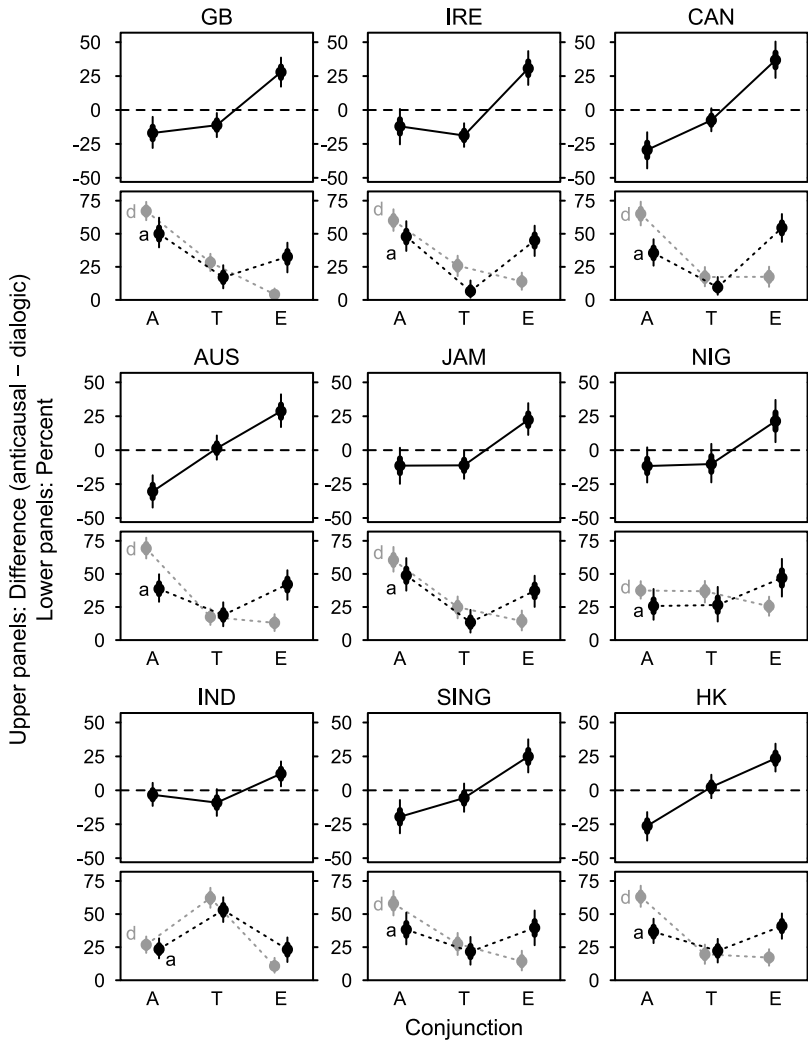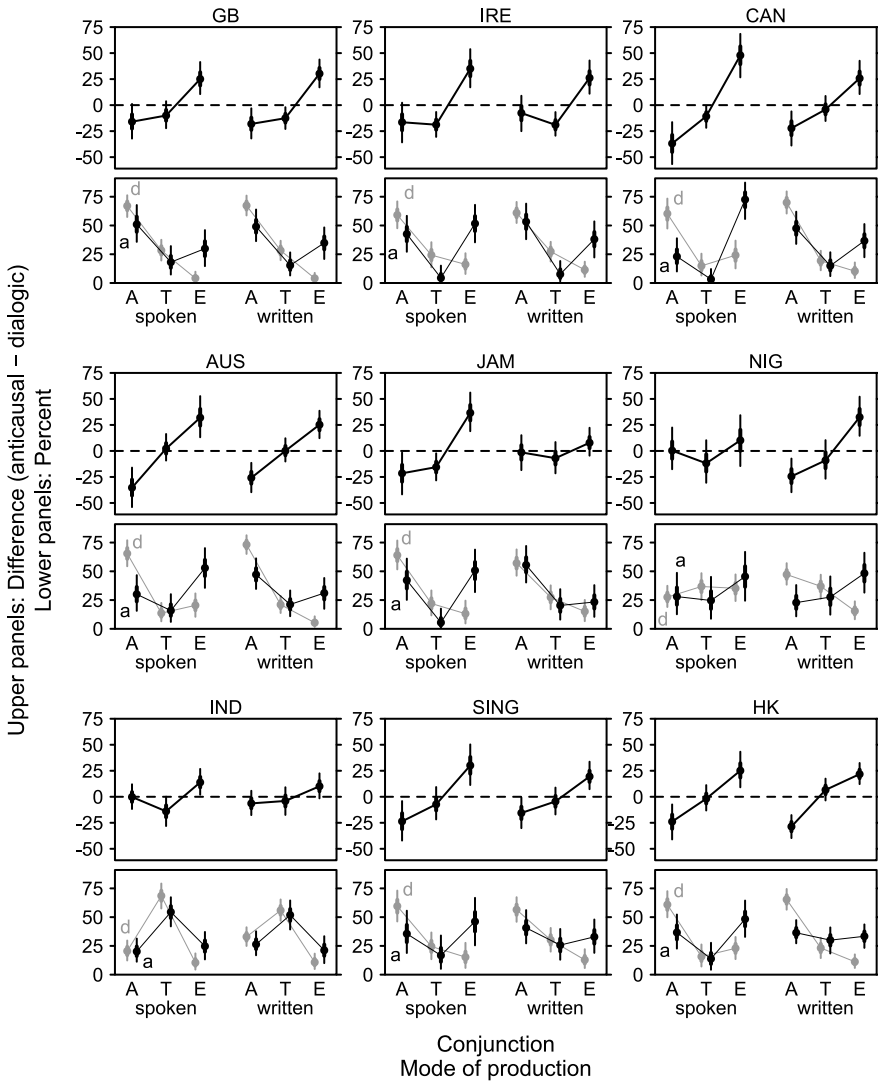
Figure 10.3: Average percentages of conjunctions by semantic type; A = *although*, T = *though*, E = *even though*, a = anticausal, d = dialogic

There is evidently a fundamental difference between anticausal and dialogic CCs concerning the roles of the conjunctions *although* and *even though*. The bird's-eye perspective fully confirms this: The mean percentage of *although* across all varieties in dialogic CCs is 56.4%, while in anticausal CCs it is 38.3%; conversely, the average percentage of *even though* in dialogic CCs is a mere 14.5%, while in anticausal CCs it is 40.2%. The conjunction *though* does not partake in this semantically conditioned variation to the same extent. As can be seen in Figure 10.3, this marker also tends to be less frequent in anticausal CCs (21.0%) as compared to dialogic CCs (28.9%), but the more modest difference between these numbers suggests that the main division of labour for the marking of specific semantic relations within a construction seems to be between *although* and *even though*. It is only in IrE and IndE that *though* is more strongly affected by semantics than *although*. Interestingly, this mirrors the results presented in §10.2.1 above, where it was found that *although* and *even though* also respond more strongly to the difference between speech and writing. While *though* tends to be functionally more similar to *although* in both dimensions of variation (mode of production and semantics), it is apparently somewhat more versatile – that is, its likelihood of occurrence does not differ as radically between conditions as is the case for the other two conjunctions.

Again, we will inspect the data for general differences between L1 and L2 varieties concerning the effect of semantic types on the selection of conjunctions. Table 10.2 provides a summary, showing for each subgroup of varieties the mean percentage of each conjunction in connection with the two semantic types, as well as the difference between these conditions.

Table 10.2: Mean percentages of conjunctions in anticausal and dialogic CCs in L1 and L2 varieties

		although	*though*	*even though*
L1	anticausal	43.0	13.0	43.5
	dialogic	65.4	22.3	12.2
	anticausal – dialogic	−22.4	−9.3	31.3
L2	anticausal	34.1	28.7	36.8
	dialogic	45.7	37.9	16.2
	anticausal – dialogic	−11.6	−9.2	20.6

Like the general effect of mode of production (cf. Table 10.2), the impact of the semantic structure of a CC on the selection of the conjunction tends to be smaller in L2 varieties than in L1 varieties. Further, and again similarly to what was shown in §10.2.1, this difference between the two subsets of varieties surfaces only with regard to *although* and *even though*, while there is no such difference in connection with *though*. If it was tentatively argued above that L2 varieties appear to be stylistically less differentiated, results in this section suggest that there is also less intra-linguistic differentiation. In other words: In L1 varieties, the semantic difference between anticausal and dialogic CCs corresponds to a more substantial formal difference (i.e. a different selection of conjunctions) than in L2 varieties.

Figure 10.4 presents the same comparison between anticausal and dialogic CCs but additionally includes the spoken-written dimension. The percentage panels at the bottom of each of the nine subplots thus contain two sets of values of the kind presented in Figure 10.3. General effects of mode and semantics as discussed earlier in this chapter can partly be traced in this plot. For instance, in several varieties the highest percentage of *even though* is found in spoken anticausal CCs, followed by written anticausal, spoken dialogic and written dialogic CCs, as in IrE, CanE, AusE, SingE and HKE. However, other varieties show that there is no perfect regularity in the ranking of constraints. This is even more clearly the case for the other two conjunctions, and we will therefore turn to a more general assessment of patterns, averaging across conditions and groups of varieties. The focus will be on the magnitude of differences in the two modes of production as displayed in the upper panels of Figure 10.4. The purely visual inspection suggests that, while in most varieties patterns in speech and writing are similar, they often appear to be more compact in writing, as in CanE, AusE, JamE, IndE and SingE, for instance. A prime example of this tendency is CanE: In both modes of production, *although* and *though* associate with dialogic CCs and *even though* associates with anticausal CCs (with the respective negative and positive values in the upper panel); in writing, however, all values are closer to zero. That is, while the general pattern is preserved, it is less extreme in writing.

General patterns and the magnitudes of semantic effects are summarised in Table 10.3, which shows for both modes of production the cross-varietal average percentages of the three conjunctions, given one or the other semantic type, as well as the mean differences between them. Thus, the table effectively sums up the interaction of semantics and mode of production. It once more illustrates the general association between semantic types and specific conjunctions: In both modes of production, percentages of *although* and *though* are higher in connection with dialogic CCs, while percentages of *even though* are higher in connec-

Figure 10.4: Average percentages of conjunctions by semantic type in speech and writing; A = *although*, T = *though*, E = *even though*, a = anticausal, d = dialogic

tion with anticausal CCs. Due to the organisation of the table, another general tendency is more difficult to detect, namely the increased percentages of both *although* and *though* in written genres and the increased percentages of *even though* in speech, regardless of semantics.

Table 10.3: Mean percentages of conjunctions in anticausal and dialogic CCs in writing and speech

		although	*though*	*even though*
Written	anticausal	42.1	23.8	33.3
	dialogic	58.9	29.9	10.7
	anticausal − dialogic	−16.8	−6.1	22.6
Spoken	anticausal	34.4	17.4	46.9
	dialogic	53.8	27.6	18.0
	anticausal − dialogic	−19.4	−10.2	28.9

A complex design that takes several intra- and extra-linguistic factors (including different global varieties) into account is bound to generate results that will not be homogeneous from all perspectives. For instance, individual varieties will diverge from general patterns, possibly due to the data quality in specific corpus components (in this case of ICE), or due to other factors unknown. Cases that do not conform to the majority pattern may provide points of departure for linguists with expert knowledge and a particular interest in the respective varieties, but they are not discussed any further in this study in order to avoid the risk of post-hoc, speculative argumentation.

10.2.3 Clause position

This section is organised in parallel to the two preceding ones. A first, general approach to the effects of clause position on the selection of conjunctions is presented in Figure 10.5, which consists of nine subplots corresponding to the varieties under investigation. The panels in the lower part of each subplot show percentages of *although*, *though* and *even though* in sentence-final subordinate clauses (in black) and clauses that are in nonfinal position (grey). The upper panels show absolute percentage-point differences between the two conditions, i.e. the subtraction of percentages in nonfinal position from percentages in final position. A dashed reference line at the value of zero ("no difference") is added to the upper panel of each subplot.
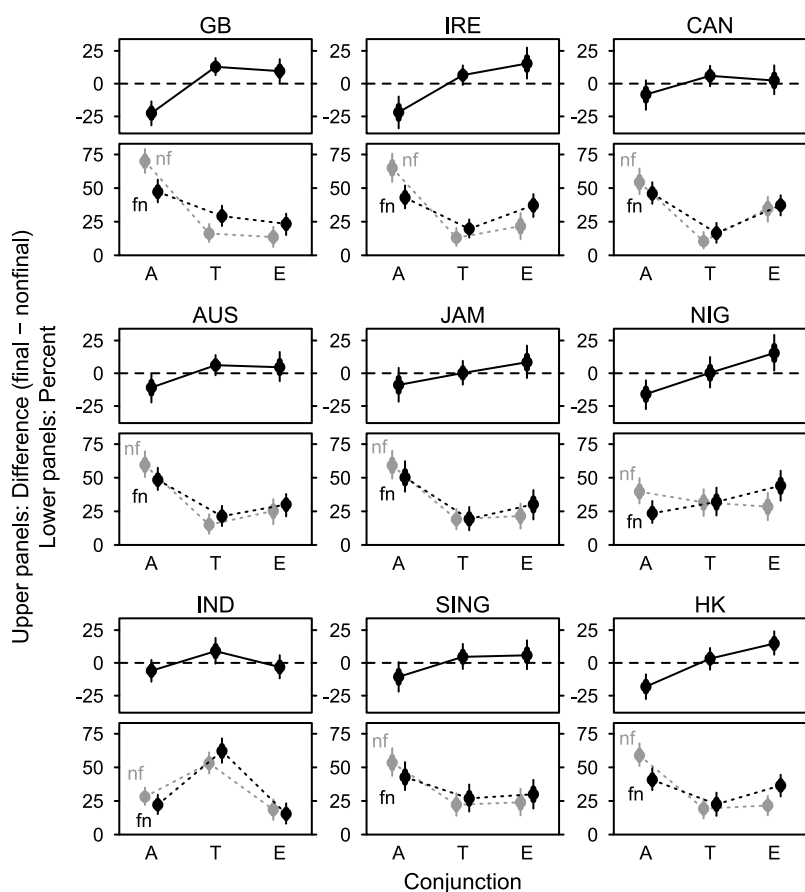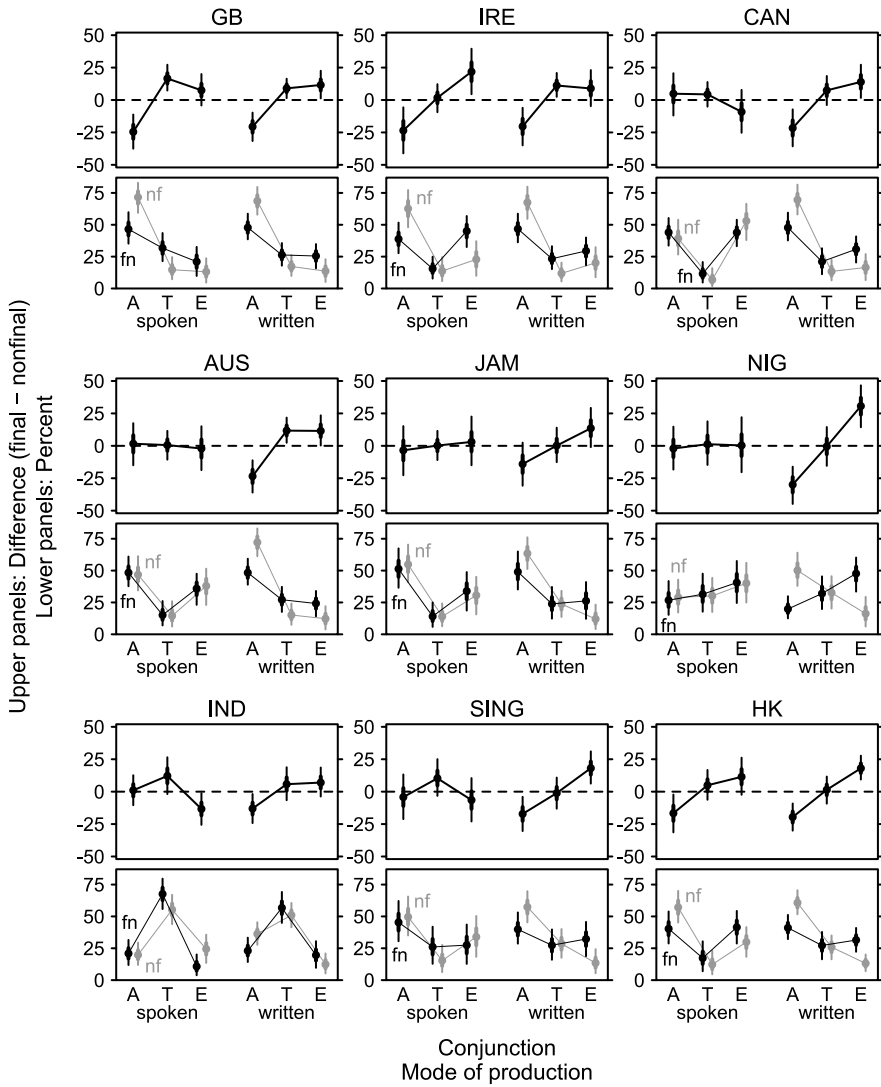
Figure 10.5: Average percentages of conjunctions by clause position;
A = *although*, T = *though*, E = *even though*, fn = final, nf = nonfinal

The main difference between the two clause arrangements concerns *although* and *even though*: Across varieties, the percentages of these two conjunctions in sentence-final clauses are 40.4% and 31.6%, respectively; in other clauses, they are 54.2% (up by 13.8 percentage points) and 23.3% (down by 8.3 percentage points). Relative frequencies of the third conjunction (*though*) are affected less (27.7% if sentence-final; otherwise 22.2%). The most common pattern for a variety across both clause positions can be described as follows: (i) In nonfinal subordinate clauses, *although* is the most commonly selected conjunction, usually followed by *even though*, with *though* coming third – albeit sometimes by a small margin; (ii) clauses in final position preserve this general pattern, with a smaller percentage-point difference between *although* and *though*. IrE, CanE,

AusE, JamE, SingE and HKE conform to this pattern. When inspecting them in Figure 10.5, we can see that, in contrast to clauses in nonfinal position, sentence-final clauses are characterised by a "flattened hockey-stick pattern", or even a V-shaped pattern.

Varieties that do not conform to this pattern are BrE, NigE and IndE. However, in BrE and NigE, sentence-final clauses are still associated with lower percentages of *although* and higher percentages of *even though*. BrE differs in using *though* more frequently than *even though* throughout, and in NigE the difference between clause positions effects a complete reversal of the frequency ranking of conjunctions. IndE simply stands out in having a unique and rather different pattern, with disproportionately high percentages of *though*. On the whole it is once again predominantly *although* and *even though* that correlate with a change in condition; *though* only shows a moderately higher percentage in sentence-final clauses.

Again, the data are inspected for general differences between L1 and L2 varieties, this time concerning the effect of clause position on the selection of conjunctions. Table 10.4 summarises for each subgroup of varieties the mean percentage of each conjunction in association with subordinate clauses in final and nonfinal position. Similarly to what was found in the inspection of mode and semantics in the previous sections, the effect tends to be smaller in L2 varieties. This is true for *although* and *though*, but not for *even though*. We can also see that the general percentage patterns in combination with each of the two clause arrangements is more level in L2 varieties – i.e. values for the three connectives are closer to each other. This is particularly visible in connection with clauses in final position, where a share of roughly one third is estimated for each of the three markers. Once again, we can carefully draw on the concept of differentiation (Schneider 2003), or a slight modification thereof: The conjunctions under investigation are possibly used less discriminately in L2 varieties, while in L1 varieties there appears to be a higher degree of specialisation, with a general preference of *although*, particularly in subordinate clauses that precede the matrix clause. This is an interesting finding because it persists in the different analyses conducted in this section. Whether or not these patterns are the result of a diachronic process of grammaticalisation and differentiation that has progressed further in L1 varieties is beyond what this study can investigate.

In Figure 10.6, the inspection of clause placement and its effects on the selection of conjunctions is unfolded into speech and writing. For each mode, there are again two sets of values in the percentage panels at the bottom of each of the nine subplots, rendered in grey (nonfinal) and black (final). Once again, the focus will lie on the direction and magnitude of differences in the two modes of

Table 10.4: Mean percentages of conjunctions by clause position in L1 and L2 varieties

		although	*though*	*even though*
L1	final	46.1	21.7	32.0
	nonfinal	62.2	13.7	23.9
	final − nonfinal	−16.1	8.0	8.1
L2	final	35.9	32.6	31.2
	nonfinal	47.8	29.1	22.9
	final − nonfinal	−11.9	3.5	8.3

production in the upper panels of the figure. Patterns are manifold and it is difficult to generalise across them. Three varieties have an indifferent or relatively flat pattern in speech that is augmented in a regular fashion in writing: JamE, NigE and HKE. The remaining four varieties (BrE, IrE, CanE and IndE), however, are not captured by this generalisation, since effects either do not differ much (or not systematically) between modes, or because the pattern is reversed, as in CanE and IndE. It would appear, then, that mode of production and clause position do not interact systematically in conditioning the selection of concessive conjunctions. The tendencies discussed above are also summarised in Table 10.5, which compares the effects of clause position on the selection of conjunctions for both modes of production, showing in each case mean percentages of clauses in final and nonfinal position as well as the difference between conditions.

Table 10.5: Mean percentages of conjunctions by clause position in writing and speech

		although	*though*	*even though*
Written	final	40.4	29.5	29.7
	nonfinal	60.6	24.4	14.4
	final − nonfinal	−20.2	5.1	15.3
Spoken	final	40.3	25.6	33.3
	nonfinal	47.8	19.6	31.8
	final − nonfinal	−7.5	6.0	1.5

The table highlights numerically what was stated above, namely that the written mode tends to augment the effect of clause position on the choice of marker.

Figure 10.6: Average proportions of conjunctions by clause position in speech and writing; A = *although*, T = *though*, E = *even though*, fn = final, nf = nonfinal

This is in contrast to the finding in §10.2.2 that semantic effects are reduced in writing. It appears that written language does not generally minimise the constraints that operate on the realisation of CCs. However, clause positions could be argued to constitute a slightly different case: While semantic properties of CCs are broadly language-internal, describing the relationship between the two propositions that make up the construction, clause positions are a formal property of CCs, and we are thus looking at a correlation of one formal parameter (clause position) with another formal parameter (choice of connective). It could be the case that particular surface forms have been codified as part of the written mode with some degree of independence from functional parameters, which would explain why (i) semantic effects are somewhat subdued in writing (see §10.2.2) and why (ii) more distinct formal realisations are brought out in writing. Of course, this interpretation is post hoc, not motivated from theory, and it therefore has to be treated with due caution.

10.2.4 Complete factor combinations

This section shows all individual conditions and their effects on the selection of markers. Sections 10.2.1–10.2.3 involved some degree of simplification (or abstraction), as the plots and discussions there were based on average values for certain conditions in which one or several of the factors were controlled for. In contrast, this section shows all the details and thus makes the underlying specific values transparent. It cannot, however, result in alternative interpretations.

The logic behind the three complex plots presented in this section is the same as in Figures 7.5–7.7, as well as in Figure 9.6, but it will nevertheless be explained in brief. There is one plot for each of the three conjunctions (*although*, *though* and *even though*), each showing the expected percentages of the respective marker in all of the $n = 72$ conditions. This number of conditions results from the fact that specific estimates differ according to variety ($n = 9$), mode of production ($\times$ 2), semantics ($\times$ 2) and clause position ($\times$ 2). For each conjunction, these conditions are made explicit on the left-hand side of the plot, and they are arranged in descending order according to their median estimates. The right-hand part of the plot highlights groupings of conditions based on variety type (L1 vs L2), mode, semantics and clause position. Once again, by comparing higher and lower concentrations of white and black squares, the reader has quicker visual access to general tendencies in the data. Each column in this part of the plot additionally shows the mean rank for each group, using triangular markers corresponding in colour to the respective group. Let us first turn to Figure 10.7, which shows the complete set of individual estimates for *although*.
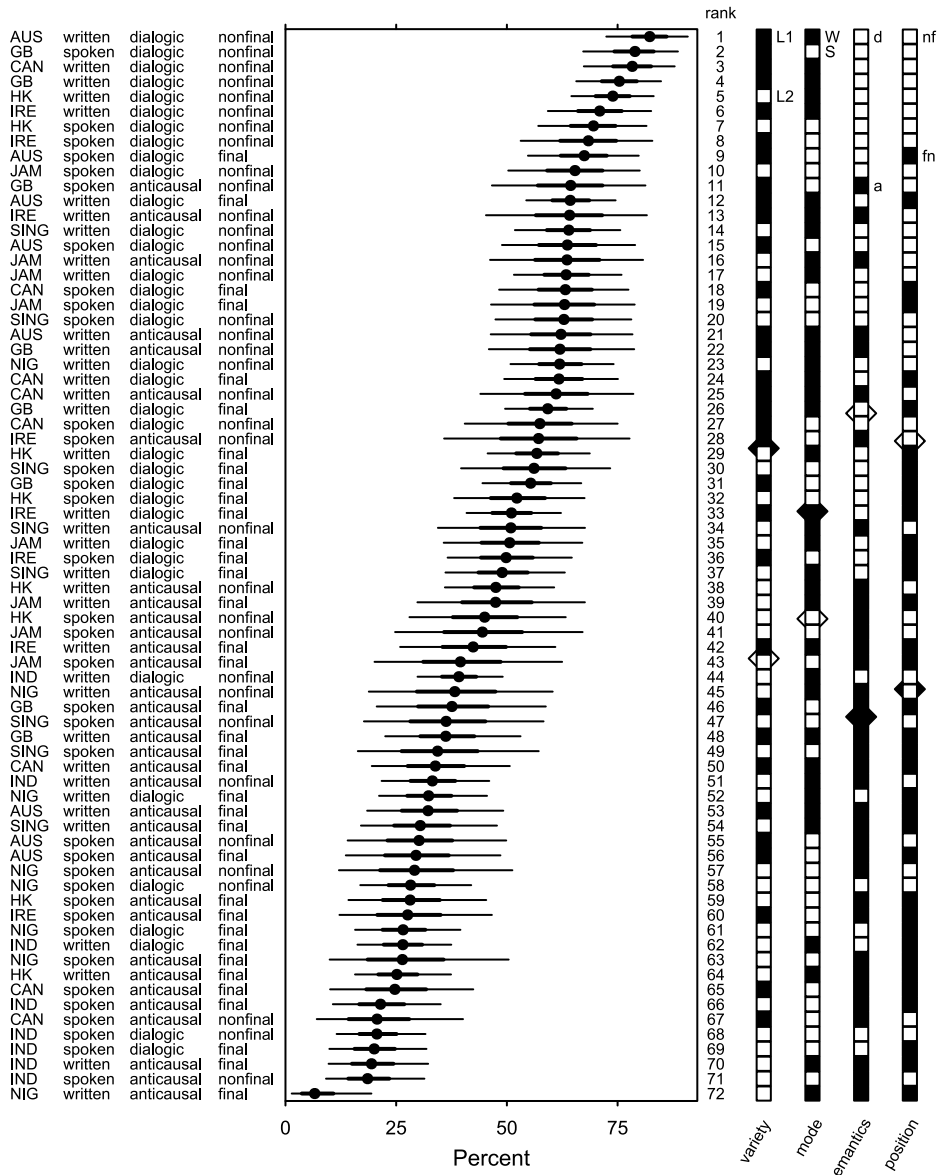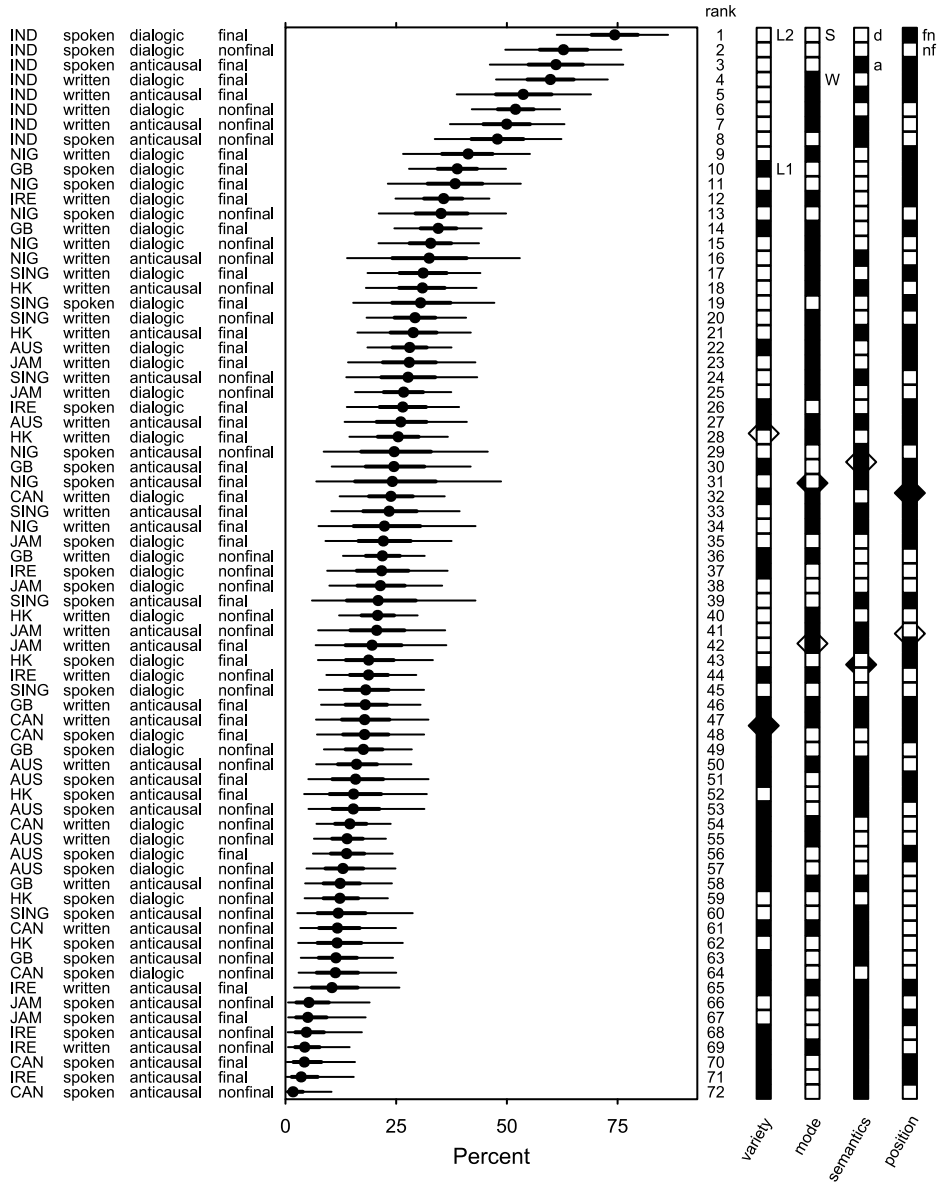
Figure 10.7: Ranked percentages of *although* by specific conditions; W = written, S = spoken, a = anticausal, d = dialogic, fn = final, nf = nonfinal

The highest percentage (rank #1) of *although* is estimated for dialogic CCs with subordinate clauses in nonfinal position in written AusE, at 82.3% [72.4; 90.9]; the lowest percentage (rank #72) is estimated for anticausal CCs that occur in sentence-final subordinate clauses in written NigE, at a mere 6.6% [1.4; 19.4]. This minimum value constitutes an outlier, but between rank #1 and rank #71 (anticausal CCs in nonfinal subclauses in spoken IndE) there is a fairly even distribution of median values. Turning to the right-hand part of Figure 10.7, we see that the distribution of black and white squares in the four columns reflects the results discussed earlier in this section. For instance, the mean rank of specific conditions from the L1 group of varieties is 28.7, while conditions in the L2 group rank considerably lower, at an average rank of 42.8; this is in line with the discussion in §10.2.1 to the effect that percentages of *although* are, on average, higher in L1 varieties. Likewise, written varieties rank higher than spoken varieties ($M_\mathrm{W} = 32.9$; $M_\mathrm{S} = 40.1$) – again, see §10.2.1 for percentage-based results that correspond to this finding. Looking at the third column, the rank-based approach illustrates that semantics (i.e. the difference between anticausal and dialogic CCs) have a larger effect than variety status and mode of production: Dialogic CCs strongly favour *although*, with a mean rank of 26.3, while the mean rank of anticausal conditions is considerably lower, at 46.7 (cf. §10.2.2). Finally, note the clear tendency for conditions to favour *although* when a subordinate clause is in nonfinal position (mean rank: 28.2) as compared to cases with clauses in final position (mean rank: 44.8) – this finding points to the "grounding" function of *although* (cf. §10.2.3). As stated above, the presentation of results in Figure 10.7 does not add substantially to the earlier discussion. It does, however, make the individual patterns transparent: While the scenarios compared in the previous three sections involved some degree of simplification since they backgrounded one or several factors, all details are shown here. Further, it is demonstrated that mean ranks as indicated in Figure 10.7 are quite reliable as a basic – and relatively intuitive – measure of effect size: The further apart the triangular indicators attached to the columns on the right, the more distinct the two basic groups that are being compared. The discussion of the conjunction *though* in Figure 10.8 happens along similar lines but will be kept somewhat shorter.

The highest-ranking estimate for *though* is for dialogic CCs with subordinate clauses in final position in spoken IndE, at a value of 74.3% [61.3; 86.4]; the lowest-ranking percentage is estimated for anticausal CCs with subordinate clauses in nonfinal position in spoken CanE, at 1.7% [0.2; 10.4]. Ranks 66–72 as well as the top eight ranks appear to break away from the central part of the distribution in Figure 10.8, which makes a somewhat more skewed impression compared to Figure 10.7 – in other words: The ordinary range of values is somewhat narrower
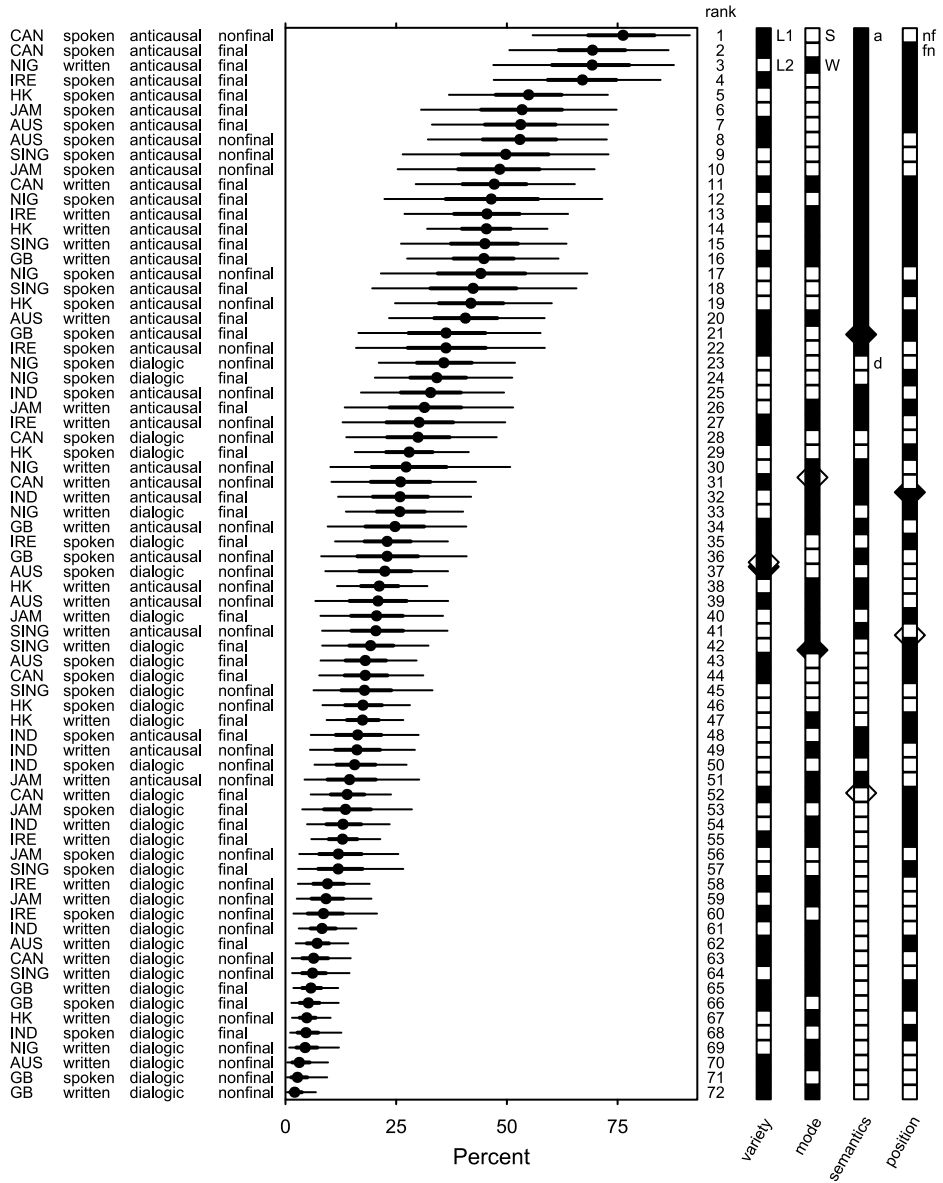
Figure 10.8: Ranked percentages of *though* by specific conditions across varieties; W = written, S = spoken, a = anticausal, d = dialogic, fn = final, nf = nonfinal

if we disregard the more extreme ranks. Once again, the distribution of black and white squares in the four columns on the right of Figure 10.8 is in accordance with earlier discussions in this section. L2 varieties (mean rank: 27.8) are considerably more likely than L1 varieties (mean rank: 47.4) to select *though* – this is of course partly due to the exceptional position of IndE, which occupies the top eight ranks. Written varieties are generally more likely to select *though* than spoken varieties, with a mean rank of 31.1 (as compared to 41.9 in speech). In the third column, we see that semantics have a similar (if somewhat weaker) effect when compared to *although* in Figure 10.7 above: The conjunction *though* is more likely in dialogic and less likely in anticausal CCs, with mean ranks of 29.7 and 43.3, respectively. The effect of clause position on the selection of *though* is the inverse of its effect on the selection of *although*, and it is also somewhat weaker: The average rank of conditions that involve subordinate clauses in nonfinal position is 41.2, while for sentence-final clauses this value is 31.8.

Let us now turn to the discussion of specific estimates for the conjunction *even though* in Figure 10.9. This marker is most frequent in anticausal CCs with subclauses in nonfinal position in spoken CanE, at 76.2% [55.8; 91.3], and its occurrence is least likely in dialogic CCs with subordinate clauses in nonfinal position in written BrE, at 2.1% [0.3; 6.9]. Apart from the top four ranks, the distribution of values is quite even. The mean ranks of specific conditions from the L1 and L2 groups of varieties in the right-hand part of the figure are virtually the same, at 36.7 and 36.4, respectively. Again, the perspective taken in this section does not generate new insights but merely shows results that were discussed earlier in a different light: Note, for instance, that the similarity of ranks for L1 and L2 varieties necessarily corresponds to the similarity of mean percentages discussed in §10.2.1 (L1: 28.0%; L2: 27.1%). In contrast to the other two conjunctions, *even though* is more likely in speech (mean rank: 30.7) than in writing (mean rank: 42.3). The strong semantic effect highlighted in the third column is also the inverse of what was found for *although* and *though*: The mean rank for conditions that involve anticausal semantics is 21.1; for dialogic semantics, the mean rank is considerably lower, at 51.9. Finally, from the perspective of ranked specific conditions, the general relationship between clause position and the selection of *even though* seems to be very similar to what was found for *though*: When a subordinate clause in nonfinal position is involved, the mean rank of conditions is 41.3 (*though*: 41.2); when there is a clause in final position, the mean rank is 31.7 (*though*: 31.8). This is in marked contrast to the rankings found for *although* with regard to this parameter.

The rank-based assessments of the four basic contrasts – according to variety type (L1 vs L2), mode of production, semantics and clause position – show from

Figure 10.9: Ranked percentages of *even though* by specific conditions across varieties; W = written, S = spoken, a = anticausal, d = dialogic, fn = final, nf = nonfinal

a different perspective the same tendencies that were discussed in §10.2.1–10.2.3. At the same time, they make the estimates for specific conditions – each defined by a unique combination of factors – maximally transparent. The conjunction *although* remains the most frequent conjunction in most scenarios. However, for a sizeable number of factor combinations it is *even though* that is estimated to be the most likely choice. These tendencies as well as the precise numbers of particular rankings are shown in Table 10.6.

Table 10.6: Frequency rankings of conjunctions relative to each other, based on *n* = 72 conditions

	Three-way ranking	*n*
1	*although > though > even though*	27
2	*even though > although > though*	17
3	*although > even though > though*	15
4	*though > although > even though*	7
5	*though > even though > although*	3
	even though > though > although	3
	Pairwise comparison	*n*
1	*although > though*	59
2	*although > even though*	49
3	*though > even though*	37
4	*even though > though*	35
5	*even though > although*	23
6	*though > although*	13
	Most frequent conjunction	*n*
1	*although*	42
2	*even though*	20
3	*though*	10

As a matter of course, a condition that is likely to produce higher percentages of one of the three conjunctions must produce lower percentages of one or both of the others. Therefore, the *n* = 72 percentages discussed above will under normal circumstances be negatively correlated for any pair of conjunctions. In Figure 10.10, these relationships are explored in some more detail, taking *although*, *though* and *even though* as the respective points of reference on the *x*-axis in the three panels of the plot. Relationships are gauged more precisely by additionally

showing the respective coefficient from a simple linear regression model, which indicates by how much the percentage on the *y*-axis changes as the percentage on the *x*-axis increases by one point.
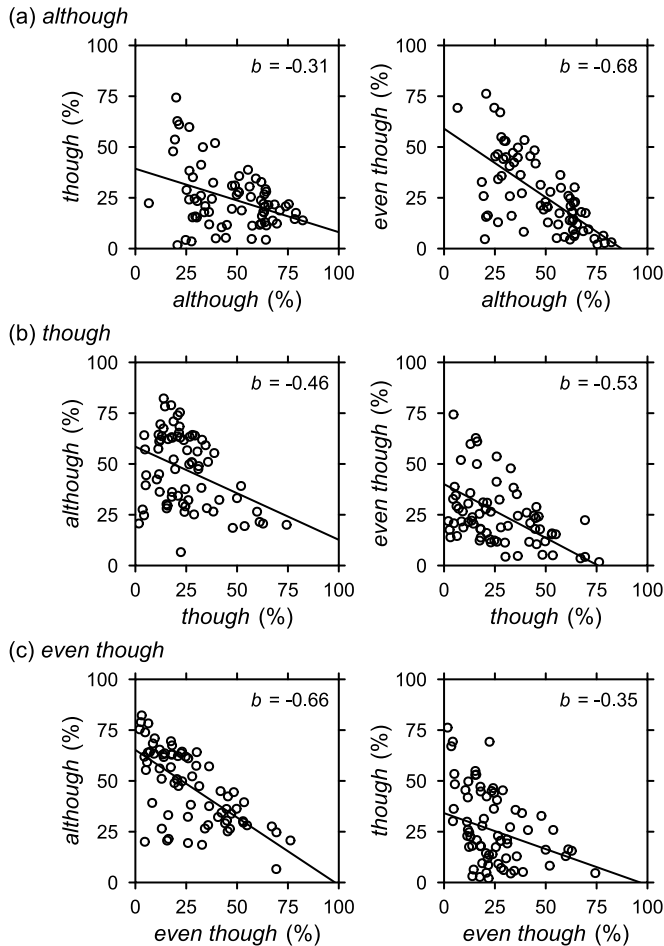
Figure 10.10: Relationship of median percentages of three conjunctions for all conditions

There are two basic, methodologically reassuring findings. Firstly, all correlations are negative. This means that any increase in the percentage of one of the conjunctions comes at the expense of both of the others – any other pattern would have been surprising in view of the results that were discussed earlier. Secondly, the beta-coefficients in each of the three parts of the figure roughly add up to one. This must necessarily be the case: The sum of percentages across all

three markers must remain at 100%, so that an increase by one percentage point in one of them must be accompanied by a total decrease of one percentage point in the other two combined. The interesting detail on whose discussion we can conclude this section is that the strongest negative correlation exists between *although* and *even though*, indicated by the regression coefficients and the steepness of the regression lines in the second plot in Figure 10.10a and the first plot in Figure 10.10c. For instance, if moving from one condition to another increases the estimated percentage of *although* by ten points, the estimated percentage of *even though* will on average decrease by 6.8 points, while for *though* the decrease will only be 3.1 points. This dovetails with the earlier discussions, in which it was found that for all four basic factors – variety type (L1/L2), mode, semantics and clause position – the greatest average swing in percentages, as we move from one factor level to the other, tends to be between *although* and *even though*. By contrast, *though* is also systematically affected but shows a more moderate response. To use once again an expression introduced in §10.2.2: The conjunction *though* is functionally more versatile, and the greatest functional contrast is between *although* and *even though*.

10.3 Summary and discussion

This final section will first summarise the main conditions likely to result in the selection of each of the three conjunctions. After thus highlighting the functional differences between the three markers, the discussion will turn to the moderating effect that mode of production has on the other main factors, as well as the general differences that were found in the comparison of L1 and L2 varieties of English in this chapter.

Under most circumstances, the conjunction *although* is the most frequent one of the three markers (see Table 10.6). It is particularly common in writing, when dialogic meaning is expressed, and when the subordinate clause is in nonfinal position. The tendency for *although* to be more frequently selected in writing is in broad agreement with the general notion that this conjunction is more formal than, for instance, *though* (Quirk et al. 1985, Biber et al. 1999, Huddleston & Pullum 2002; also Aarts 1988). Further, the association of *although* with dialogic semantics agrees with patterns in AmE data in Schützler (2018a) as well as NZE data in Schützler (2017), but it is contra the general tendencies described in Hilpert (2013a).[1] The clear difference between the present study and Hilpert

---

[1]Schützler (2017) includes data from the British, Canadian and New Zealand components of ICE; the first two data sources can of course not be cited as additional (independent) evidence, since they also feature in the present study.

(2013a) is surprising only at first glance: Due to his particular research interest in concessive parentheticals, Hilpert focuses on CCs with co-referential subjects in both clauses, which narrows the eligible constructions down to a grammatically more restricted and therefore smaller set. The strong association of dialogic CCs with *although* can also account for the high text frequency of this marker (cf. Chapter 7): Since the dialogic type is the most common kind of CC (as shown in Chapter 8), and since dialogic semantics tend to be expressed with *although*, the high overall frequency of this conjunction naturally follows.

When we average across different conditions and thus gloss over the differences induced by the various internal and external factors, the conjunction *though* seems to be of roughly the same relative frequency as *even though*. However, like *although* (and in contrast to *even though*) it associates mostly with dialogic CCs, which strengthens it in terms of text frequency (cf. Chapter 7). The probability of selecting *though* is higher in writing, but this effect is usually weaker than for *although*. This finding casts doubt on the assertion that *though* is a less formal variant, which is found in some of the literature (particularly in the major standard grammars; but see also Aarts 1988). If we accept writing and speech as very basic stylistic categories, we would expect less formal items to occur at higher frequencies in spoken discourse. This is not the case for *though*. All we can say is that this conjunction is affected less strongly by a difference in mode than *although*, but both effects are in the same direction. In contrast to *although*, however, the conjunction *though* tends to be used more in subordinate clauses in final position and resembles *even though* in this respect.

Lastly, *even though* is considerably more frequent in speech, in contrast to the other two conjunctions. The literature has very little to say about this marker's formality value but regularly stresses its emphatic character, presumably triggered by the adverb *even*. It could be argued that emphasis and immediacy are more characteristic of speech, in the sense that SP/W draws on material that is felt to be stronger or more emotive in order to persuade AD/R. This higher degree of emphasis coincides with the fact that *even though* is morphologically the most complex of the three markers (cf. §2.3). Once again in contrast to the other two conjunctions, *even though* is strongly associated with anticausal CCs. This finding explains why this marker appears to be quite rare from a perspective purely based on text frequency; in the variationist approach, i.e. when we consider variable contexts and the factors that play a role in the selection of conjunctions, there are many scenarios (particularly in anticausal CCs) in which *even though* can be quite frequent, or even the most frequent variant. In marked contrast to *although* but to some extent similar to *though*, *even though* is more common in subordinate clauses that follow the matrix clause.

In the more general comparison of the three markers, it is striking that *although* and *even though* seem to form the poles of a functionally motivated (probabilistic) continuum: *although* is associated with written discourse, dialogic semantics and subordinate clauses in nonfinal position; *even though* is associated with spoken discourse, anticausal semantics and subordinate clauses in final position. This goes hand in hand with the finding that these two markers respond more strongly to differences in mode of production and semantics, compared to *though*.

As regards the differences between L1 and L2 varieties of English, the sample in the present study is of course too small for sweeping generalisations ($n_{L1} = 4$; $n_{L2} = 5$). Thus, the tendencies that were detected need to be treated with caution and can be taken as no more than indicators with the potential of providing guidance for future research. In the L1 varieties, the average effect of mode of production on the selection of subordinators is greater than in the L2 varieties. The same is true with regard to the effect of semantics. A tentative conclusion that agrees with a broad understanding of Schneider's (2007) notion of differentiation in Phase 5 of his Dynamic Model is the following: Patterns of use in L2 varieties are somewhat more fixed, i.e. they respond less sensitively to conditioning factors, be they external (e.g. mode, or style more generally) or internal (e.g. semantic or information-structural). In other words, varieties from this broad subset have undergone less formal differentiation along contextual and functional lines. The systematic variability of rules (to use a key concept from variationist linguistics) is equally visible and tends to be in the same direction in L1 and L2 varieties, but effects tend to be smaller in the latter.

For the final part of this summary, I will return to the differences and similarities between the three conjunctions. It was argued that the main division of labour is between *although* and *even though*, while *though* tends to be functionally more intermediate: Typically, *although* is used for grounding purposes (putting the matrix clause in focus position at the sentence level), for dialogic CCs and for written discourse; *even though* associates with anticausal CCs in which the subordinate clause is in final (i.e. focus) position, and it is more common in speech; *though*, like *although*, is more typical of writing and dialogic CCs, but – like *even though* – it tends to be attached to subordinate clauses that follow the matrix clause. Seeing that there is a functional continuum with *though* at its centre, and that the three markers can (and regularly do) serve exactly the same purposes, it is unsurprising that the literature thus far has either treated them as functionally equivalent or has tried to capture differences exclusively in terms of categories like "emphasis" or "formality". However, as this chapter has shown, we can profile the differences between the three conjunctions in a

more nuanced way and demonstrate that they are not only measurably different but also form a system in which specific tasks are assigned to specific markers. Naturally, those tendencies are not categorical but probabilistic. At present, we can only speculate as to why it is *although* and *even though* that are (or have become) particularly specialised in several respects. Ultimately, the answer to this question needs to be sought in diachronic studies on a similar scale as the present synchronic one, i.e. studies based on data sets large enough to include the same (or perhaps even more) factors. A new, diachronic hypothesis generated from the present research would be that, over time, the morphological variants *although* and *even though* grammaticalised into functionally somewhat different items. The pattern I would expect to find in diachronic data is therefore one of gradual functional divergence. On the one hand, we might see *although* and *even though* slowly breaking away in different directions, increasingly specialising on the marking of constructional variants diametrically opposed in terms of typical contexts of use (e.g. mode), semantics and general syntactic design (clause sequencing). On the other hand, we would expect that *though* does not undergo the same degree of specialisation but borrows characteristics from the other two connectives, because in PDE it is more likely in final position, in writing and in combination with dialogic CCs. Filling in the diachronic details of such processes, or investigating whether or not such processes can be shown to have taken place at all, goes far beyond what the present study can achieve. I will return to these thoughts and their implications for future research as part of the final discussion in Chapter 12.

11 Clause structure

Analysing the internal structure of subordinate clauses in CCs is the final step when progressing through the stages of the choice model formulated in §4.1.3. The realisation of a subordinate clause as finite or nonfinite depends on all other factors – semantic structure, clause position and the connective itself. In analogy to Chapters 7–10, this chapter first presents the statistical model that was used (§11.1), shows the results of the analysis (§11.2) and discusses them against the expectations that were formulated (§11.3).

11.1 Statistical model

The model employed for the analysis of clause-internal syntax ("Model E") is a binary logistic regression model since the outcome variable NONFIN takes only two values, "nonfinite" and "finite". The latter is the reference category and also happens to be the unmarked, much more frequent variant overall (see Table 6.3 in §6.3.6 for an overview of variables). The model, shown in (87), is essentially constructed in the same way as Models C & D (see §9.1 and §10.1): Mode of production (represented by the variable SPOKEN.CT) interacts with each of the other fixed-part predictors, but these do not interact with each other. Additionally, the full set of fixed-part predictors, with the exception of SPOKEN.CT, are assumed to vary randomly across the cluster variables GENRE and TEXT. As in the previous analyses, separate models with identical specifications were run for each of the nine varieties.

(87) Model E: Syntax

```
nonfin ~ spoken.ct * (anti.ct + final.ct + marker)
         + (anti.ct + final.ct + marker | genre)
         + (anti.ct + final.ct + marker | text)
```

Although it contains the largest number of predictors, Model E is in fact intermediate in complexity between the less complex Model C and the more complex Model D, since the latter is a multinomial model that generates a considerably larger number of parameters. More information concerning token numbers, the

number of levels of both random factors (GENRE and TEXT), the priors (held constant across all nine models) and the number of posterior samples can be found in Appendix B.5. Data, scripts and model summaries (i.e. tables with regression coefficients) can be retrieved from the online repositories (cf. §1.4).

11.2 Results

This part of the chapter takes five perspectives on the results to highlight different factors and their impact on the outcome. Firstly, the effects of semantics, clause position and the concessive conjunction are controlled for in §11.2.1. This results in average values for individual varieties, and the only variety-internal differentiation happens along the spoken-written dimension. Next, §11.2.2 isolates the effect of intra-constructional semantics, again considering the moderating effect of mode of production. Thirdly, in §11.2.3 the focus will lie on the relationship between clause position and the finite/nonfinite status of subordinate clauses. In the fourth section (§11.2.4), likely combinations of conjunctions and nonfinite clauses are explored. Finally, §11.2.5 shows a complete ranking of specific conditions, in analogy to the approach taken in §10.2.4.

11.2.1 Average percentages

As indicated above, the first step in analysing relative frequencies of finite and nonfinite subordinate clauses was to establish average values for the nine varieties. Again, the assumption of neutral values for mode of production, semantics and clause position is a hypothetical one, but it helps to arrive at a general impression. Figure 11.1 shows percentages of nonfinite clauses in CCs for the nine varieties under investigation, arranged in ascending order. Once again, L1 varieties are shown in black and L2 varieties in grey.
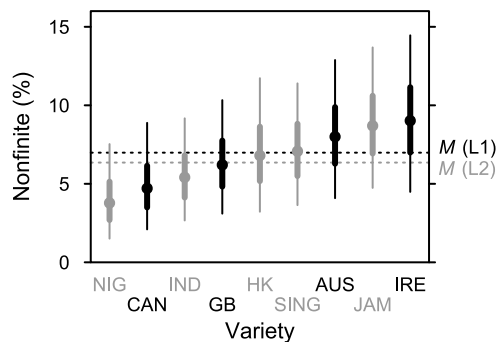
Figure 11.1: Average percentages of nonfinite subordinate clauses

A relatively small number of subordinate clauses are realised as nonfinite, the range of variety-specific median values extending from 3.8% [1.5; 7.5] in NigE to 9.0% [4.5; 14.5] in IrE. There is only a small difference between L1 and L2 varieties: The mean percentage of nonfinite clauses in the former is 7.0% (as indicated by the dotted black line), while in the latter it is 6.4% (as indicated by the grey line). Given the great variability between varieties in combination with the relatively high degree of intra-varietal uncertainty, this is not a substantial difference, and it will accordingly not be discussed any further.

Figure 11.2 again orders varieties according to their average percentages of nonfinite subordinate clauses. This time, however, results in the lower panel are subdivided into values for the spoken and written mode (in grey and black, respectively). Additionally, the estimated difference between modes of production (in absolute percentage points) is shown in the upper panel. Mean differences between speech and writing are indicated separately for L1 and L2 varieties, using dotted lines with direct labels.
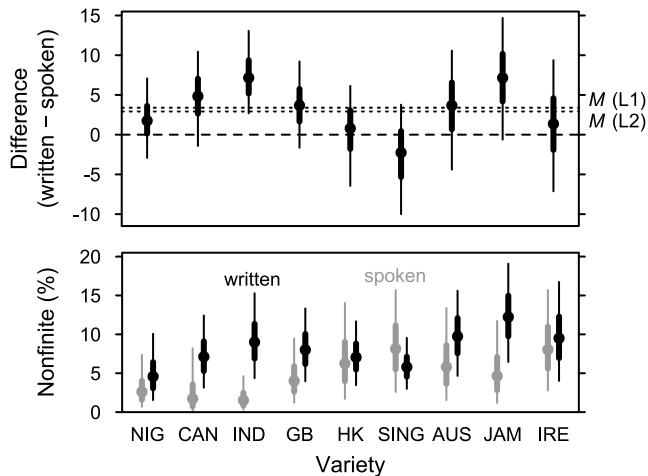
Figure 11.2: Average percentages of nonfinite subordinate clauses by mode of production

The share of nonfinite subordinate clauses tends to be higher in writing. This general pattern is found in eight out of nine varieties, with SingE as the only exception. Five varieties have a somewhat more substantial positive percentage-point difference (AusE: +3.7; BrE: +3.7; CanE: +4.8; JamE: +7.2; IndE: +7.2), but only IndE has a value that seems robustly different from zero, with a 90% uncertainty interval of [2.7; 13.1]. If the mean difference between writing and speech is inspected separately for L1 and L2 varieties, it turns out to be only minimally

larger among the former (+3.4) compared to the latter (+2.9). Table 11.1 explores potential differences between L1 and L2 varieties in some more detail by providing separate mean values for the spoken and written mode. In this table, differences between writing and speech in L1 and L2 varieties are not exactly the same as in the upper panel of Figure 11.2: A slight discrepancy arises between the per-group means of estimated, variety-specific differences on the one hand (as in Figure 11.2) and the difference between group-specific mean estimates for percentages on the other (as in Table 11.1; cf. comment in Footnote 4 on p. 141), and rounding errors may also differ. To avoid confusion, further comparisons of this kind will therefore only be based on values shown in tables.

Table 11.1: Nonfinite realisations of subordinate clauses by mode and variety type (mean %)

	spoken	written	written − spoken
L2	4.6	7.7	3.1
L1	4.9	8.6	3.7
L2 − L1	−0.3	−0.9	

The small general difference between L2 and L1 varieties, with the former characterised by slightly fewer nonfinite subordinate clauses, is remarkably similar in both modes of production. From the bird's-eye perspective shown in Figure 11.1 above (i.e. controlling for mode of production), this difference was 0.5 absolute percentage points. In speech and writing, it is 0.3 and 0.9 percentage points, respectively. There is neither a substantial difference between the two broad groups of varieties concerning finiteness/nonfiniteness, nor do the two groups differ markedly in their response to a change in mode of production.

11.2.2 Semantics

The lower part of Figure 11.3 isolates the effect of the intra-constructional semantics of CCs on the finiteness of subordinate clauses, with dialogic and anticausal types shown in grey and black, respectively. In the upper panel, the differences between the two conditions are shown, this time without an indication of mean values for L1 and L2 varieties. The plot applies the same horizontal ranking as the previous two figures, based on the average percentages of nonfinite clauses in the individual varieties.
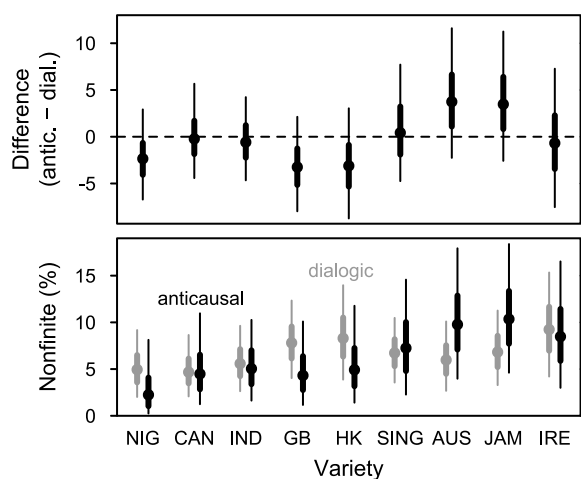
Figure 11.3: Average percentages of nonfinite subordinate clauses by variety and semantic type

The relationship between intra-constructional semantics and the finiteness status of the subordinate clause seems highly unsystematic – in fact, the general pattern of differences in the upper panel of Figure 11.3 is remarkably similar to the respective panel in Figure 11.2 above. Four varieties (CanE, IndE, SingE and IrE) hardly make a difference between the two conditions, three varieties (NigE, BrE and HKE) tend to use fewer nonfinite clauses in dialogic CCs, and the remaining two varieties (AusE and JamE) have a higher percentage of nonfinite clauses in dialogic CCs. In all cases, the difference comes with high degrees of uncertainty. Table 11.2 provides a few basic summary statistics that underscore the absence of a pattern along this dimension of variation: Only individual varieties stand out from a rather nondescript general distribution, and there is little that could be said about differences between L1 and L2 varieties.

Table 11.2: Nonfinite realisations of subordinate clauses by semantics and variety type (mean %)

	anticausal	dialogic	anticausal − dialogic
L2	6.0	6.5	−0.5
L1	6.8	6.9	−0.1
L2 − L1	−0.8	−0.4	

Unfolding the global patterns outlined above into the spoken and written mode, as shown in Figure 11.4, adds relatively little to our understanding of the relationship between semantics and (non)finiteness. As in the analogous plots in previous chapters, there is one component plot per variety. Estimated percentages of nonfinite clauses by mode and semantics are shown in the lower panels, while absolute percentage-point differences between anticausal and dialogic CCs – still distinguishing the two modes – are displayed in the upper panels. In contrast to the plots above, varieties are no longer ordered on quantitative grounds but according to the original arrangement introduced in §4.3 and §6.1 (cf. Table 4.1 and Figure 6.1).

There are only few patterns that suggest a moderating effect of the mode of production on the selection of a nonfinite or finite subordinate clause. In most cases, the effect of semantics on finiteness is either extremely (CanE, NigE and IndE) or very (BrE, AusE and HKE) similar between modes; in the remaining three varieties (IrE, JamE and SingE), there is a more substantial difference between modes, but it lacks coherence in that the moderating effect points in different directions.

These findings do not suggest that there is a systematic, readily interpretable connection between the internal structure (finiteness/nonfiniteness) of a subordinate clause and the semantic relation that holds between clauses within a CC. It is quite probable that the few sporadic, variety-specific patterns that we *can* see are due to the fact that each variety was addressed with its own separate model. Had the approach been to fit a single model, with VARIETY as a grouping variable and the semantic predictor ANTI.CT varying randomly across it, the pooling effect would quite possibly have reduced the effect even further (cf. §6.3.1), particularly seeing the low and evenly distributed absolute numbers of nonfinite cases as documented in Appendix A.3.

11.2.3 Clause position

In analogy to the approach in the previous section, the lower panel of Figure 11.5 directly contrasts in a simplified form the effect that clause positions have on the realisation of a subordinate clause as finite or nonfinite. Values representing subclauses in nonfinal position are shown in grey, while values for clauses in final position appear in black.[1] The upper panel shows the differences that result when subtracting percentages (of nonfinite constructions) in nonfinal positions

---

[1] It is awkward that the terms *final/nonfinal* and *finite/nonfinite* are so similar, phonologically. To increase processability, only *finite* and *nonfinite* will be used as direct attributes of *clause*; that is, I will speak of *finite/nonfinite clauses* and *final/nonfinal positions*, but not of "final/nonfinal clauses".
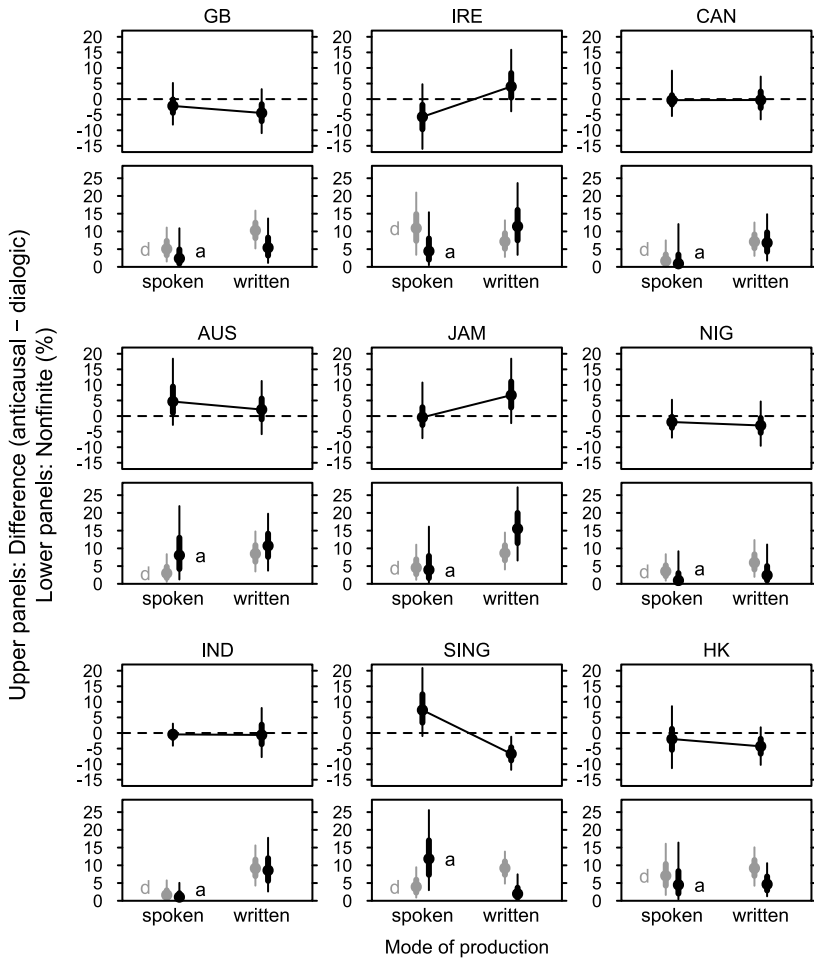
Figure 11.4: Average percentages of nonfinite subordinate clauses by variety, mode and semantic type; a = anticausal, d = dialogic

from those in final positions. The same arrangement of varieties as in the similar plots in §11.2.1 and §11.2.2 is retained.
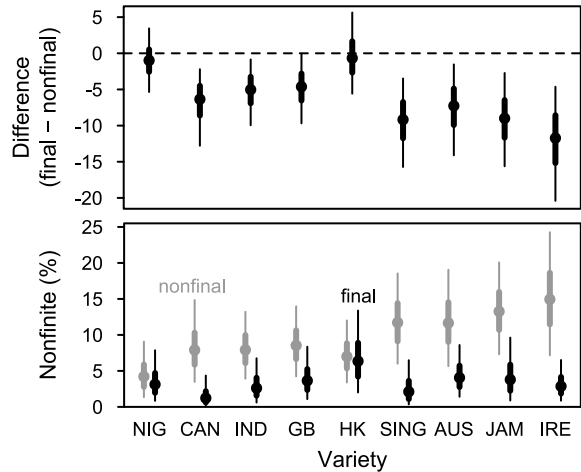
Figure 11.5: Average percentages of nonfinite subordinate clauses by variety and clause position

Compared to the rather noisy semantic pattern in the previous section, there is a much more systematic relationship between clause position and the internal structure of subordinate clauses: The proportion of nonfinite realisations is always lower if the subordinate clause is in final position. Two of these differences are very close to zero (NigE: −1.0 [−5.4; 3.4]; HKE: −0.7 [−5.6; 5.6]), but the remaining seven are not, with BrE (−4.6 [−9.7; −0.1]) and IrE (−11.7 [−20.4; −4.6]) forming the extreme points of the group. A few basic summary statistics are produced in Table 11.3, comparing L1 and L2 varieties. Given the discussion above, the basic pattern – with higher percentages in nonfinal positions – must of course obtain in both subsets, but the contrast between conditions is somewhat more striking in L1 varieties.

Table 11.3: Nonfinite realisations of subordinate clauses by clause position and variety type (mean %)

	final	nonfinal	final − nonfinal
L2	3.6	8.8	−5.2
L1	2.9	10.8	−7.9
L2 − L1	0.7	−2.0	

The interaction of mode of production and clause position in conditioning the selection of (non)finite clause realisations is shown in Figure 11.6. Three varieties – IrE, AusE and SingE – show a relatively level pattern in the upper panel, which means that the difference (in absolute percentage points) between final and nonfinal clause positions is fairly stable across modes of production. In this group, the intuitively most plausible pattern can be seen in AusE, where both values (for final and nonfinal positions) are lower in speech. In IrE, the percentage is relatively stable across modes for clauses in nonfinal position, while for clauses in final position it is lower in speech. The SingE pattern is somewhat puzzling since values for both clause positions are slightly higher in speech, which goes directly against the general trend and the expectation that was formulated – compare Figure 11.2, in which SingE was the only variety characterised by a higher percentage of nonfinite clauses in speech.

Apart from the level pattern (with more or less horizontal lines in the upper panels of Figure 11.6) described above, several varieties show a very clear interaction effect, whereby percentages of nonfinite realisations in combination with nonfinal clause positions are pulled towards zero in speech – that is, the environment normally favouring nonfinite realisations (that is, nonfinal position) does not do so in this mode of production, and the expected percentages become much more similar to those associated with clauses in final position. In the plots of difference in the upper panels of Figure 11.6, the resultant pattern is one with the difference in speech close to zero and a relatively steep downward slope when moving to the right-hand part of the plot, representing writing. This can be seen in CanE, JamE, NigE and IndE.

Finally, in BrE and HKE there is a "crossed" pattern: In speech, it is predominantly clauses in final position that take nonfinite complements, while in writing it is clauses in nonfinal positions that do. There is thus a positive difference (% final −% nonfinal) in speech and a negative one in writing. This finding does not conform to the formulated expectations. Given the rather low overall token numbers for nonfinite realisations (see Appendix A.3), we can only speculate as to the reasons, which probably lie in sampling errors or confounding factors to do with the specific information structure of the few instances that are involved.

In sum, the analyses in this section suggest that a nonfinite clause realisation is normally more likely if the subordinate clause is in nonfinal position. This is not unexpected, as it aligns well with Quirk et al.'s (1985) principle of "communicative dynamism", whereby heavier and more informative syntactic elements tend to occur later in a sentence (cf. §2.3.1). Since nonfinite clauses not only lack a finite verb but very often also a subject, they have less material substance and their early placement therefore comes as no surprise. The matrix clause, on the
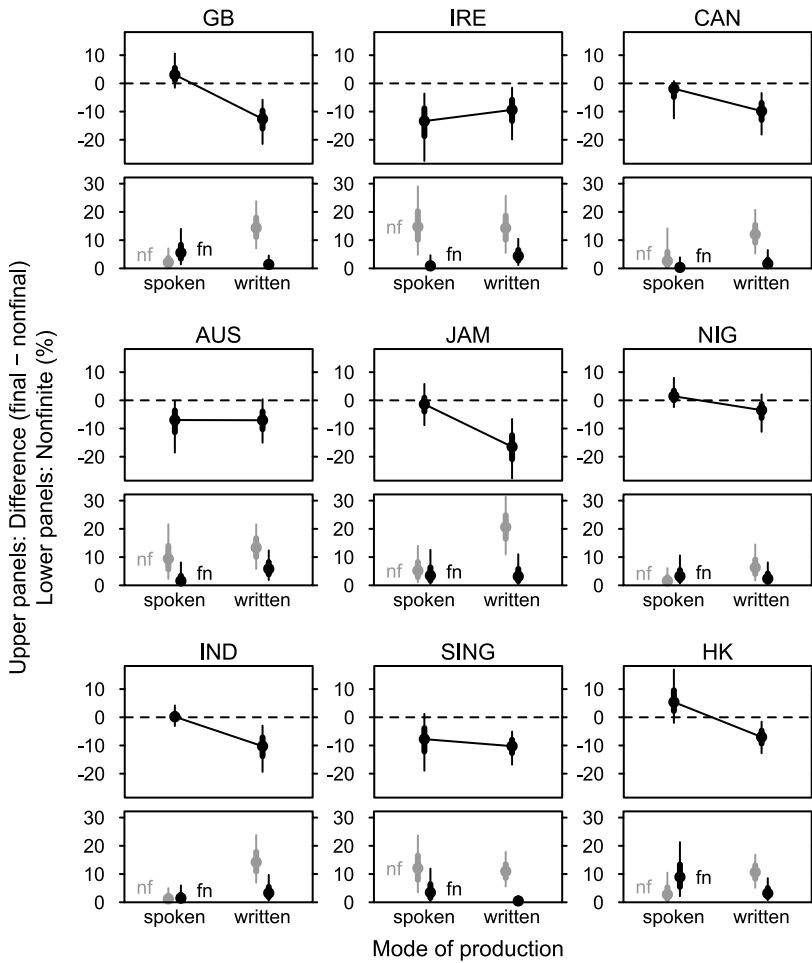
Figure 11.6: Average percentages of nonfinite subordinate clauses by variety, mode and clause position; fn = final, nf = nonfinal

other hand, has both a finite verb and a syntactic subject and on average tends to be placed after a nonfinite subordinate clause.

11.2.4 Markers

In this section, the focus is on the relationships between the three concessive conjunctions *although*, *though* and *even though* and the realisation of a subordinate clause as finite or nonfinite. Because in this case three conditions are compared, Figure 11.7 differs in design from the respective first plots in §11.2.1–11.2.3, and it does not show estimates of differences.[2]
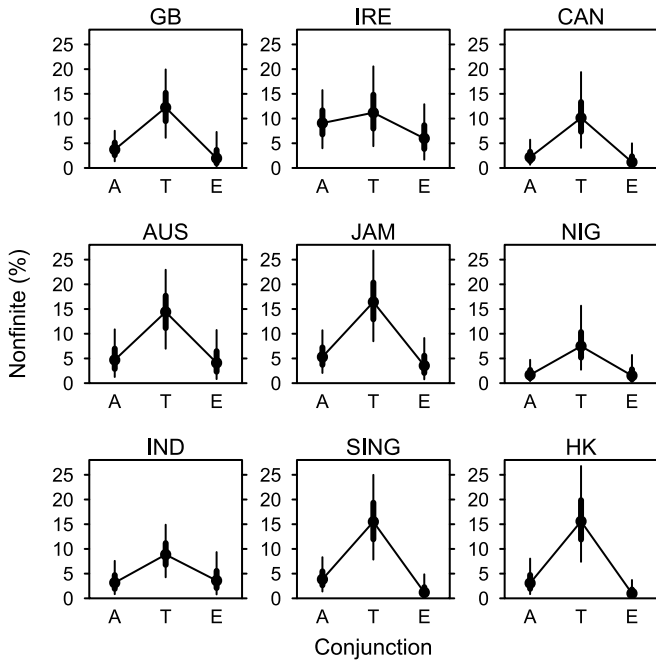
Figure 11.7: Average percentages of nonfinite subordinate clauses by variety and marker; A = *although*, T = *though*, E = *even though*

---

[2]Pairwise differences between markers could have been estimated, but their interpretation would have been challenging.

Although the slopes of lines connecting the three values in the individual panels differ in steepness, the general similarity of patterns is very striking. In eight varieties, the expected percentages follow a uniform ranking, namely *though > although > even though*. Averaging across the median estimates of all nine varieties, we get mean values of 12.4% for *though*, 4.1% for *although* and 2.7% for *even though*. IndE constitutes the single exception to this pattern, with percentages of 3.6 for *even though* and 3.2 for *although*. However, the difference between *although* and *even though* is generally not very large, and the inverted ranking in IndE is not very striking. The affinity between *though* and nonfinite clauses is in accordance with Hilpert's (2013a) findings (cf. §5.1.4). The highest overall value is estimated for *though* in JamE (16.4% [8.5; 26.8]); the lowest value is found for *even though* in HKE (1.0% [0.2; 3.7]). Looking back to the previous sections of this chapter, it appears that the rather low overall percentage of nonfinite subordinate clauses in the global perspective – with a cross-varietal mean value smaller than 7% (cf. Figure 11.1) – results from the fact that in those parts of the analysis the effects of the markers themselves were neutralised.

The inspection of average values according to the two broad groups of L1 and L2 varieties in Table 11.4 does not reveal any substantial difference between them concerning the correlation of (non)finiteness and specific markers. The general ranking described above (*though > although > even though*) holds within both groups, and the absolute percentage-point difference between L1 and L2 varieties for each individual marker does not seem remarkable, either.

Table 11.4: Nonfinite realisations of subordinate clauses by marker and variety type (mean %)

	although	*though*	*even though*
L2	3.4	12.8	2.2
L1	4.9	12.0	3.3
L2 − L1	−1.5	0.8	−1.1

The analysis next turns to the interaction of mode of production and individual conjunctions in conditioning the selection of (non)finite subordinate clause realisations. The focus in Figure 11.8 is on the absolute percentage-point difference between writing and speech, as shown in the respective upper panels of the nine subplots. Values above the dashed reference line indicate that nonfinite subordinate clauses are more frequent in writing (which is the expected pattern), while values below the line signify that they are more common in speech.
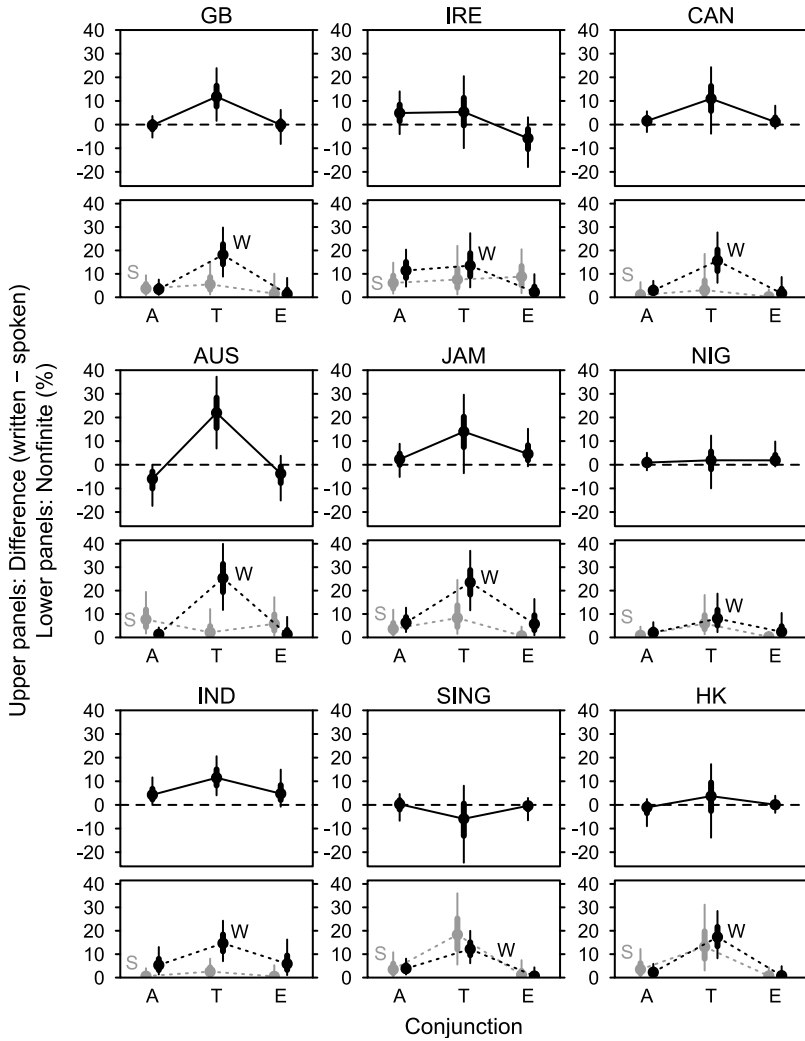
Figure 11.8: Average percentages of nonfinite subordinate clauses by variety, mode and marker; A = *although*, T = *though*, E = *even though*, W = written, S = spoken

Once again, there are relatively clear tendencies, although in comparison to Figure 11.7 the number of exceptions is somewhat larger. Most differences (% written − % spoken) are in a positive direction or close to zero, which confirms that written language is characterised by a higher share of nonfinite subordinate clauses in CCs. The only exceptions (i.e. tendencies in the opposite direction) sufficiently different from zero to deserve discussion are *even though* in IrE (−5.8 [−18.0; 3.1]), *although* in AusE (−6.0 [−17.4; −0.3]), and, with some reservations, *though* in SingE (−5.9 [−24.4; 8.1]). Typically, *though* is the marker that responds most strongly to the difference in mode of production, as reflected in the positive wedge-shaped patterns in the upper panels of Figure 11.8 for BrE, CanE, AusE, JamE, IndE and HKE. In this set, AusE and HKE display the most and the least pronounced patterns of this kind, respectively. Concerning the remaining three varieties, there are conflicting (i.e. hard-to-interpret) patterns in IrE and SingE, and a level pattern in NigE. The strong affinity between *though* and nonfinite subordinate clauses suggests that this particular combination of formal characteristics qualifies as a subconstruction, and this view is further supported by its particular sensitivity to differences in mode of production.

The ranking of conjunctions according to their co-occurrence with nonfinite subordinate clauses is also interesting at a more general level. The shortest conjunction (*though*) is most likely to introduce nonfinite subordinate clauses, which will on average also be relatively short, due to the absence of a finite verb and a grammatical subject. Conversely, the longest conjunction (*even though*) is the one most likely to combine with finite – and therefore longer – subordinate clauses, followed by the second longest marker, *although*. Thus, in terms of the weight of subordinate clauses, we effectively get a split into (i) longer constructions that combine complex/long markers with syntactically unreduced/finite clauses and (ii) shorter ones that combine the marker *though* with reduced/nonfinite clauses. It could be argued that the special function of *though* aids AD/R in parsing the sentence, since the occurrence of this particular marker signals an increased likelihood of a following nonfinite (and therefore cognitively somewhat more complex) clause. However, despite the tendencies shown in this section, it is of course still the case that in combination with *any* of the three conjunctions finite clauses remain in the majority. As will be shown in the next section, this is true even if all factors are set against nonfiniteness.

11.2.5 Complete factor combinations

This section provides the final, most detailed perspective on the estimated share of nonfinite subordinate clauses expected to occur under different circumstances.

In contrast to the previous sections, no averaging across specific conditions is applied but all possible combinations of factor settings are shown. Their total number is *n* = 216 (9 varieties × 2 modes of production × 2 semantics × 2 clause positions × 3 markers). Due to the large number of conditions, ranked estimates are shown in three consecutive plots: Figure 11.9 shows ranks 1–72, Figure 11.10 shows ranks 73–144, and Figure 11.11 shows ranks 145–216. The percentage scale once again has a horizontal orientation.[3] To the right of each figure, the scheme of grey-scale symbols known from earlier chapters (cf. §7.3, §9.2.2 and §10.2.4) is used to highlight structure in the data. In the first four columns, black squares represent L1 varieties, written language, anticausal semantics and subordinate clauses in final position, respectively; conversely, white squares denote L2 varieties, spoken language, dialogic semantics and subordinate clauses in nonfinal position. The fifth column differentiates between the three subordinators *although*, *though* and *even though*, using black, grey and white boxes, respectively. Additionally, the mean ranks of groups of conditions – "black" vs "white" (vs "grey") – are indicated in each column by the triangular markers known from earlier plots. Those average ranks (like the ranks themselves) are established based on all three figures in combination. Since this way of plotting the data merely reveals the underlying specifics and does not add novel insights to the analysis, the discussion in this section will be kept relatively brief. Note that, in contrast to earlier plots, only 50% uncertainty intervals are shown.

The first rank is occupied by dialogic CCs in written BrE whose subordinate clauses are in nonfinal position and headed by *though*, with an expected percentage of nonfinite clauses of 43.6%, directly followed by anticausal CCs in written JamE with subordinate clauses in nonfinal position introduced by *though* (43.4%). The models predict a number of relatively high values (at the top of Figure 11.9), but only *n* = 36 of the estimated *n* = 216 specific median percentages are actually above 10% (i.e. exactly one in six conditions). Ranks 165–216 (*n* = 52, which corresponds to 24% of all cases) round to a whole-number value of zero on the percentage scale, as indicated by the dashed horizontal line in Figure 11.11. Turning to the mean ranks calculated for sets of conditions grouped according to basic predictor values, we necessarily obtain patterns that support the findings documented in §11.2.1–11.2.4, as discussed in the following paragraph. Most indicators of ranks for such groups (that is, the triangular markers added to the columns on the right of the three plots) are found in Figure 11.10, i.e. among the middle

---

[3]The numerous low percentage values are very difficult to discriminate in Figure 11.10 and Figure 11.11. An alternative way of plotting percentages using logit scaling and thus increasing the resolution of low (and high) values is discussed in Schützler (2023) but not applied here.

third of ranks; only the mean ranks for conditions involving *though* and *even though* are found in Figures 11.9 & 11.11, due to the strong association of these conjunctions with finite and nonfinite clause realisations, respectively.

The very slight – not to say, negligible – general difference between the two broad variety types (L1 vs L2) that was discussed above (cf. Figure 11.1) corresponds to the only marginally higher mean rank of conditions involving L1 varieties (M = 105.7) compared to conditions involving L2 varieties (M = 110.7). The pattern of black and white squares in the first of the analytic columns in the three figures above does not reveal any obvious structure. Contrasting written and spoken varieties in the second column, there is a clearer pattern: Particularly if we compare Figure 11.9 to Figure 11.11, we observe a greater density of black squares in the former (among the top 72 ranks) and a greater density of white squares in the latter (among the bottom 72 ranks). The mean ranks are M = 96.3 for writing and M = 120.7 for speech, which agrees with the general patterns observed in Figure 11.2 above. A considerably smaller difference is once again found for groups based on semantics, as shown in the third column in the three figures. The mean rank for dialogic CCs is 105.8, while for anticausal CCs it is 111.2 – the difference between the two semantic types was even difficult to perceive in Figure 11.3 above (see also Table 11.2). On the other hand, a very substantial difference was found in the general comparison of subordinate clauses in final and nonfinal position (see Figure 11.5), and the three figures in this section reflect this, with mean ranks of 131.1 for the former and 85.9 for the latter. Finally, the associations of the three concessive conjunctions with the (non)finiteness of subordinate clauses can be traced in the fifth column of Figures 11.9–11.11. In §11.2.4, *though* emerged as the marker correlating most strongly with nonfinite subordinate clauses, while *although* and particularly *even though* are much less likely to introduce such clauses (see Figure 11.7 above). Figures 11.9–11.11 throw these patterns into relief: The average rank of conditions involving *even though* is 146.0, for *although* it is 115.9, and for *though* it is 63.6, which is the only value to make it into the top 72 ranks.

The presentation of fully specified conditions in this section has naturally confirmed the more general scenarios discussed in the earlier parts of the chapter. However, the inspection of all n = 216 possible factor combinations provides a more realistic impression of how those values were arrived at, namely by averaging across a large number of low percentages – very often close to zero – and a small number of higher values. Under most circumstances, the realisation of a subordinate clause as nonfinite remains the exception: As Figures 11.9–11.11 show, written discourse, nonfinal clause position and the conjunction *though* need to coincide to generate a more substantial share of this particular syntactic type.
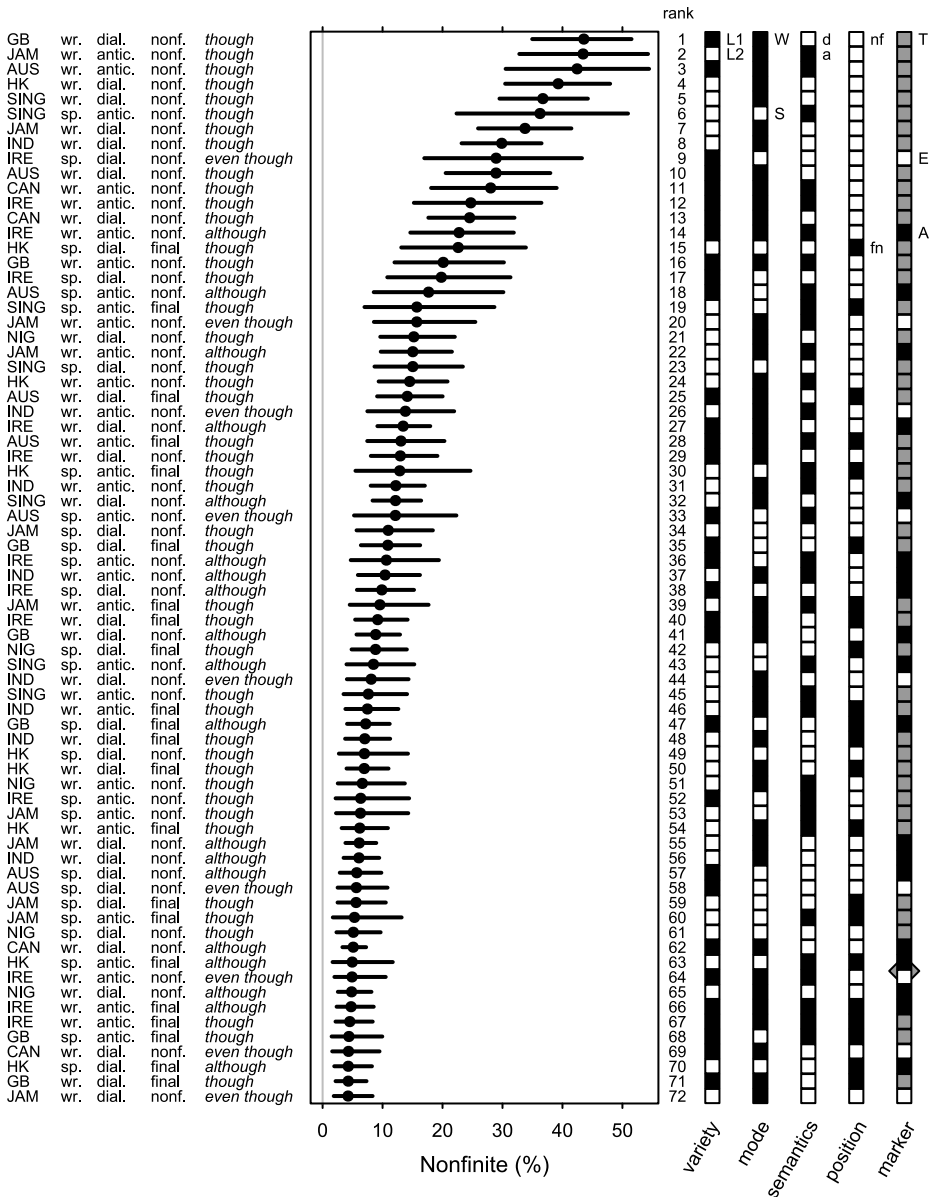
Figure 11.9: Ranked percentages of nonfinite clauses by specific conditions, ranks 1–72; with 50% uncertainty intervals; W = written, S = spoken, a = anticausal, d = dialogic, fn = final, nf = nonfinal, A = *although*, T = *though*, E = *even though*

Figure 11.10: Ranked percentages of nonfinite clauses by specific conditions, ranks 73–144; with 50% uncertainty intervals; W = written, S = spoken, a = anticausal, d = dialogic, fn = final, nf = nonfinal, A = *although*, T = *though*, E = *even though*

variety	mode	semantics	position	marker	rank
AUS	sp.	antic.	final	though	145
IRE	sp.	dial.	final	though	146
IND	sp.	dial.	final	even though	147
NIG	wr.	antic.	final	even though	148
CAN	sp.	dial.	final	though	149
JAM	sp.	dial.	nonf.	even though	150
IRE	sp.	dial.	final	although	151
AUS	wr.	dial.	nonf.	even though	152
IND	sp.	dial.	final	although	153
IRE	sp.	antic.	final	though	154
IRE	sp.	antic.	final	although	155
AUS	wr.	dial.	final	although	156
SING	wr.	antic.	final	though	157
NIG	sp.	dial.	final	even though	158
GB	sp.	antic.	nonf.	even though	159
CAN	sp.	antic.	nonf.	although	160
HK	sp.	antic.	final	although	161
AUS	sp.	dial.	final	although	162
IND	sp.	dial.	final	although	163
CAN	wr.	dial.	final	although	164
IRE	sp.	antic.	final	even though	165
AUS	sp.	dial.	final	even though	166
JAM	wr.	dial.	final	although	167
JAM	sp.	antic.	nonf.	even though	168
AUS	wr.	antic.	final	even though	169
JAM	sp.	dial.	final	even though	170
HK	wr.	antic.	nonf.	even though	171
IRE	sp.	dial.	final	even though	172
HK	wr.	dial.	final	even though	173
IND	wr.	dial.	final	although	174
IND	sp.	antic.	nonf.	although	175
JAM	sp.	antic.	final	even though	176
NIG	sp.	dial.	nonf.	although	177
IND	sp.	dial.	final	even though	178
IND	wr.	dial.	final	even though	179
CAN	sp.	antic.	final	though	180
IRE	wr.	dial.	final	even though	181
CAN	sp.	dial.	nonf.	even though	182
AUS	sp.	dial.	final	though	183
SING	wr.	antic.	nonf.	even though	184
IND	wr.	antic.	final	even though	185
HK	sp.	dial.	nonf.	even though	186
IND	sp.	antic.	nonf.	even though	187
CAN	sp.	dial.	nonf.	even though	188
NIG	wr.	antic.	final	although	189
SING	sp.	dial.	final	although	190
SING	wr.	dial.	final	although	191
GB	wr.	antic.	final	although	192
HK	sp.	antic.	nonf.	even though	193
HK	sp.	antic.	final	even though	194
SING	sp.	antic.	final	even though	195
CAN	wr.	antic.	final	even though	196
GB	wr.	dial.	final	although	197
CAN	wr.	dial.	final	even though	198
IND	sp.	dial.	nonf.	even though	199
NIG	sp.	dial.	nonf.	even though	200
JAM	wr.	dial.	final	even though	201
GB	wr.	antic.	final	even though	202
IND	sp.	dial.	nonf.	although	203
NIG	sp.	antic.	final	although	204
SING	sp.	dial.	final	even though	205
SING	wr.	dial.	final	even though	206
NIG	sp.	antic.	nonf.	although	207
CAN	sp.	dial.	final	although	208
GB	wr.	dial.	final	even though	209
SING	wr.	antic.	final	although	210
NIG	sp.	antic.	final	even though	211
CAN	sp.	antic.	final	although	212
CAN	sp.	dial.	final	even though	213
SING	wr.	antic.	final	even though	214
NIG	sp.	antic.	nonf.	even though	215
CAN	sp.	antic.	final	even though	216

Values below line round to zero.

Nonfinite (%): 0 10 20 30 40 50

Legend: L1, L2, S, W, a, d, fn, nf, T, E, A

Figure 11.11: Ranked percentages of nonfinite clauses by specific conditions, ranks 145–216; with 50% uncertainty intervals; W = written, S = spoken, a = anticausal, d = dialogic, fn = final, nf = nonfinal, A = *although*, T = *though*, E = *even though*

11.3 Summary and discussion

The main factors that play a role in the selection of nonfinite and finite subordinate clauses in CCs are (i) mode of production, (ii) clause position and (iii) the subordinating conjunction. There is neither a systematic difference between L1 and L2 varieties, nor do the intra-constructional semantics of a CC seem to have an impact on the internal structure of subclauses.

Concerning mode of production, results support the hypothesis formulated in §5.3: Less explicit (elliptical) nonfinite subordinate clauses are more common in writing, arguably because the challenges they pose for the processor are alleviated in this mode. The association of nonfinite structures with written discourse is well-known from the literature. It is a typical feature of a more compressed style, and therefore requires no additional discussion here.

No hypotheses were formulated concerning the relationship between clause position and the (non)finiteness of subordinate clauses. I have argued that nonfinite subordinate clauses preceding the matrix clause should be more problematic from a processing perspective: They not only lack a finite verb but usually also a subject (cf. §2.3.2), so that their full interpretation must be suspended at least until the matrix clause subject is parsed. On the other hand, according to Quirk et al.'s (1985: 1036) notion of *resolution* (i.e. end-weight applied at the sentence level), the heavier clause would be expected at the end of a sentence, and this will typically be the finite matrix clause. There were thus diametrically opposed predictions, whose relative importance can only be established empirically. Results in this study suggest a strong alignment of the nonfinal placement of a subordinate clause and its realisation as nonfinite. It appears that the weight of component clauses plays a more important role than the challenge presented by a suspended subject.

Finally, the finding that nonfinite subordinate clauses are more likely to be attached to the conjunction *though* is in agreement with expectations – expectations, however, that are based exclusively on findings by Hilpert (2013a), not on theoretical considerations. Like the association of *although* with subordinate clauses in nonfinal position (see Chapter 9), the association of *though* with nonfinite clauses makes the range of possible specific constructions somewhat tidier – in concrete terms, it makes the combination of formal characteristics less arbitrary. This higher degree of orderliness in itself can motivate constructional patterns – even without invoking additional semantic, formal or language-external factors – and a more principled account of these thoughts will be provided in the final chapter of this volume. However, a few remarks on the combination of *though* with nonfinite subordinate clauses are nevertheless in place, if only to

provide pointers for future research. It is curious, for instance, that the shortest marker (*though*) should be the one that most readily combines with nonfinite (and therefore shorter) clauses, and, conversely, that the longest marker (*even though*) should most readily combine with finite (and therefore longer) clauses. Although no more than an informed speculation, there appears to be a tendency for the economy of clauses at the sentence level not to strive towards balanced constructions (short marker + long clause; long/complex marker + short clause) but to favour a somewhat more obvious differentiation into subordinate clauses that are either heavier or lighter on both counts. It would of course be interesting (and perhaps necessary) to see the emergence of such tendencies in diachrony, and thus to shed light on a specific kind of constructional change. It seems quite possible that language users actively exploit different degrees of clause weight to emphasise certain parts of sentences and thus to generate specific information structures. Secondly, *though* is historically primary, while *although* and *even though* are somewhat later additions to this set of conjunctions. Thus, there appears to be an attraction between the potentially most grammaticalised marker and types of subordinate clauses that are cognitively more complex since they contain less explicit information. Although we cannot truly derive such theories from the present research, the final chapter will point towards some possible avenues for future research that may incorporate assumptions of this kind.

12 Conclusion and outlook

This study set out to generate insights concerning the functional and formal variation of a certain set of concessive constructions (CCs), namely complex sentences with the subordinating conjunctions *although*, *though* or *even though*. The original point of departure (as in Schützler 2018b) was the question as to why these three markers coexist in English, and what the division of labour between them is. This book goes some way beyond this original question: It is no longer only the connectives that are under scrutiny, but the general correlations that exist between functional and formal properties of the constructions in which they occur. Highlighting the ties between these different facets of a CC, the study fills some of the gaps that are left by grammars of English.

The book also proposes one particular approach to constructional variation, essentially dealing with the questions of how to build theories for complex, multifaceted constructions and their variability, and how to capture those constructions in statistical terms.[1] The resulting kind of quantitative Construction Grammar treats the different components of a CC (and, by extension, other constructions) as hierarchically ordered and embedded within each other. This may seem to be in conflict with some of the basic tenets of CxG – for instance, the fusion (or inextricability) of form and function. On the other hand, it can be argued that the model agrees well with the notion that more general constructions break down into subconstructions at different levels of granularity. This paradoxical situation – with the scientific model partly supporting, partly contradicting CxG-based thinking – will be discussed in some more detail in §12.3.

Apart from its contributions to the description of CCs and to CxG-based theories, the present study is also relevant in the context of varieties of English world-wide (see §4.3). However, general results in this dimension of variation suggest that the phenomenon at hand is not a salient marker of variety affiliation, as most of the inter-varietal differences that do exist are relatively slight or unsystematic, particularly when inspecting the general contrast between L1

---

[1]Apart from their introduction in §6.3, statistical techniques were not foregrounded in the analytic chapters of this volume. The online appendix (https://osf.io/m4tfc/) – perhaps together with the published data (Schützler 2021; https://doi.org/10.18710/1JMFVR) – provides much more detail and can serve as a point of departure for further analyses (see also §1.4).

and L2 varieties.[2] On the whole, CCs and their structured variation seem to be a relatively stable and homogeneous part of English grammar, at least from the synchronic perspective.

Beyond all of the above, the present study has provided, categorised and discussed in detail a wealth of corpus examples, highlighting the semantic and pragmatic versatility of CCs. The tension between the propositions juxtaposed in a construction can be based on generally understood pieces of world knowledge (so-called *topoi*) concerning facts that are typically incompatible, yielding what was called *anticausal* concessives or their inverse, *epistemic* concessives. In *dialogic* concessives, on the other hand, the contrast may be based not on the expected incompatibility of facts but merely on the qualification of one proposition by another. Hilpert (2013a: 166) calls concessives of this kind "mixed-messages", because they allow for different overall interpretations or evaluations and may therefore trigger different, perhaps even diametrically opposed, courses of action. For all types of CCs – anticausal, epistemic and dialogic – the number of possible topoi and semantic patterns is vast, and the possible propositional content of CCs is virtually limitless. In a way, what is produced by a collection of CCs and the precise relations holding between their component propositions is essentially a mirror image of human reasoning and argumentation.

The paragraphs above can stand as a broad summary of the main contributions of this book. The remainder of this chapter serves three purposes: (i) It summarises the main results of the quantitative analyses (§12.1); (ii) it points to wider contexts of investigation in which we can place CCs, and more comprehensive ways of looking at these constructions (§12.2), venturing recommendations as well as warnings; and (iii) it reflects in more detail upon the advantages and disadvantages – as well as the overall plausibility – of the choice model of constructional variation that was proposed (§12.3), including the discussion of alternative views. Finally, §12.4 concludes the book with a few final remarks.

12.1 Summary of results

Throughout the book, a distinction was made between Chapters 7 & 8 on the one hand and Chapters 9, 10 & 11 on the other. The two earlier chapters work with the text frequencies of conjunctions and semantic types (using count models), while the three later chapters inspect choices in variable contexts (using binary and multinomial regression models). As will be discussed in §12.1.1 below, the

---

[2]Many of the patterns are perhaps best regarded as reflections of sampling error, the diachronic dimension of ICE, or differences between the individual compilation processes.

former type of analysis is somewhat limited compared to the latter: We may be in a position to explain much of what determines choices made in the relevant contexts, but it is more difficult to explain the frequency of a phenomenon as a whole. For instance, the text frequency of a particular semantic type may well depend on the discourse topic and other factors not normally of interest in a variationist approach. On the other hand, text frequency has often been central in corpus-linguistic studies, and its discussion in this book – particularly vis-à-vis the contributions made by Chapters 9, 10 & 11 – can highlight certain methodological issues. Results from the latter three chapters are discussed in §12.1.2, drawing on the choice model of constructional variation that was proposed, and thus forming a more integrated whole. Finally, §12.1.3 returns to a question that originally inspired the investigation as a whole (cf. Schützler 2018b). This concerns the functional differences between the three conjunctions *although*, *though* and *even though*, i.e. the question as to how exactly they divide between them the task of introducing concessive subordinate clauses.

12.1.1 Frequency-based accounts: Uses and limitations

As discussed in §7.4, investigations of the text frequencies of phenomena have traditionally taken centre stage in quantitative corpus linguistics. However, they may come with the risk of presenting an oversimplified picture. Absolute (normalised) frequencies sometimes do, but often enough do *not* give us the answers we are looking for. The summary of results from Chapters 7 & 8 will therefore be brief, and it will to an extent serve the purpose of throwing the discussion of variable contexts in §12.1.2 into sharper relief.

In an inspection of the cumulated frequencies of all three conjunctions it turned out that the total number of subordinating CCs was reasonably similar in most varieties, but also that rather extreme outliers do exist, e.g. BrE with its much higher overall frequency, and NigE with a very much lower overall rate. While, in a multifactorial design, the researcher can with some success discuss the reasons why one of several possible realisations was selected, general frequency differences of this kind may result from data quality issues, or from hard-to-gauge characteristics of varieties and their underlying cultures. They are therefore difficult to interpret. Concerning the individual conjunctions, *although* is usually most frequent in written English, while *even though* is most of the time least frequent. In speech, *even though* is more frequent relative to the others, mostly because it is much less susceptible to the tendency of spoken language to generate fewer complex sentences. The general pattern in the present study agrees with much of the literature (e.g. Quirk et al. 1985, Altenberg 1986, Aarts

1988), although it does not seem plausible to describe *though* as less formal than *although* (cf. Quirk et al. 1985: 1097–1099, Biber et al. 1999, Huddleston & Pullum 2002), at least not on the basis of a simplistic operationalisation of style as "spoken vs written". There is, however, a tendency for *although* to respond most strongly and for *even though* to respond least strongly to differences in mode of production, all of which aligns relatively well with patterns found by Altenberg (1986) and Aarts (1988), for instance. Quirk et al.'s (1985) characterisation of *even though* as emphatic is difficult to confirm, unless we stretch our definition of *emphasis* to simply include notions like "involvement" or "directness", which would then partly account for this marker's popularity in speech. The meaning of the results described here will be brought out more clearly by considering not only what we know about the currency of semantic types (see following paragraph) but also by the multifactorial investigation summarised in §12.1.2.

As regards the text frequencies of semantic types, the present study somewhat surprisingly found that dialogic CCs are by far the most frequent type in all varieties – "surprisingly", because grammars primarily tend to cite anticausal examples and the literature seems to treat these as prototypical. However, the anticausal type only comes second in frequency, followed at a considerable distance by epistemic (and narrow-scope dialogic) CCs. Contrary to expectation, dialogic CCs do not associate with speech. This correlation was hypothesised because the two component propositions in this semantic type are pragmatically on a par (i.e. not captured by an IF → THEN relation), the entire construction is therefore (cognitively) more coordinated in character, and paratactic structures are generally more common in speech. The finding that narrow-scope CCs are considerably more frequent in writing casts further doubt on the usefulness of text frequencies as outcomes. It can be shown that, at the syntactic level, this particular semantic type is most commonly constructed with a nonfinite complement of the conjunction. Since nonfiniteness is generally characteristic of writing, the large number of narrow-scope CCs found in that mode may in fact be an artefact of this particular syntactic property, which considerably complicates the interpretation of results.

Much of the literature remains silent on the issue of semantic types of CCs. In direct contrast with findings in the present study, results reported by Hilpert (2013a) suggest that the anticausal type is most frequent. As argued in §5.1.2, Hilpert's study does not include the conjunction *even though* and moreover focuses on specific constructions with co-referential subjects in matrix clause and subordinate clause. Particularly the second point can probably account for much of the discrepancy between results. Concerning the present study, the fact that

the predominance of dialogic CCs holds true in all varieties under investigation inspires a certain degree of confidence in this finding.

12.1.2 Multifactorial analyses at different levels

The notion of the different "levels" of a construction and its theoretical implications will feature more prominently in §12.3 below. Here, suffice it to remind the reader that the quantitative analyses of Chapters 9, 10 & 11 assume a nestedness of lower-level (or more local) constructional properties within higher-level (or more general) properties. The highest formal level involves the placement of the basic building blocks in a CC, matrix clause and subordinate clause. The intermediate level involves the selection of a concessive conjunction, which serves as the node between the two clauses. At the lowest level, the subordinate structure is syntactically unfolded into a finite or nonfinite clause. It is in this order that results will be summarised in the following paragraphs.

Clause position was treated as a binary variable, taking the values "final" and "nonfinal", with the latter comprising initial and medial positions (see §2.3.1 and §6.3.6). The main results from the analysis of variable clause positions in CCs are summarised in the following three points. A more detailed summary and discussion follows below.

1. Sentence-final position of subordinate clauses is more common in L1 varieties than in L2 varieties;

2. sentence-final position is more likely in speech; and

3. there is no systematic general link between the intra-constructional semantic relation (here: anticausal vs dialogic) and the arrangement of clauses.

Concerning the first result, it was initially hypothesised that L2 varieties would favour subordinate clauses in final position. This was based on the view that the final placement of subordinate structures is cognitively optimal, both in terms of production and parsing (cf. §2.3.1). Due to the somewhat less central and secure status of English in L2 varieties, it was argued, the cognitively less complex (perhaps: more natural) patterns would tend to prevail. I suggested that a potential reason for the unexpected inverse pattern may lie in the generally more scholastic acquisition of English in L2 contexts and the predominance of prototypical cases of anticausal CCs with preposed subordinate clauses in such settings (as foregrounded in grammar books, for instance). However, post-hoc

speculations of this type can only be substantiated with an independent research effort and will not be pursued any further here.

The finding that subordinate clauses are more likely to be placed in final position in the spoken mode agrees with the respective hypothesis, even if the effect is generally not large and nonfinal placement remains the majority variant in many spoken varieties. From the perspectives of both production and processing, final placement was considered cognitively less demanding than nonfinal placement (again, see §2.3.1), and mechanisms of this kind should of course be all the more effective in speech, due to its transient nature.

The absence of a systematic relationship between intra-constructional semantics and clause position undermines the hypothesis that clausal arrangements in anticausal CCs should be iconic of the semantic relation between propositions (once more, see §2.3.1 and §5.3). This hypothesis seemed particularly appealing, as its confirmation would have provided a plausible link between functional and formal parameters internal to the construction. However, we see an unsystematic array of patterns across varieties, some supporting, some undermining the hypothesis. In combination with the relatively weak effects for mode of production, they leave us with an uneasy feeling regarding clause position as an outcome variable. It was argued that important – and perhaps central – factors were not taken into consideration in this study. These could include the discourse-structuring intentions of SP/W, who may have a certain theme-rheme (or topic-comment) structure in mind. Thus, structures and the ways in which they present and foreground information have their motivation in the wider discourse context and in SP/W's construal of it. Since the present study treated CCs as hermetic (i.e. restricted to exactly two component clauses and their relation), other, possibly central factors must necessarily slip the net of the analysis. These issues and their implications for future research will be discussed further in §12.2.2.

Results for the *choice of conjunction* can be summarised in five points, one of them addressing the general picture, the other four commenting on specific factors and their impact on the probability of occurrence of each of the three markers. Note that in this discussion we are still moving through the hierarchy imposed by the choice model. An alternative, more holistic perspective on markers is provided in §12.1.3.

1. Controlling for individual factors, *although* is generally the most frequent marker; frequencies of *though* and *even though* are similar to each other, but much lower.

2. *Although* is selected particularly in the written mode, to express dialogic meaning, and if the subordinate clause is in nonfinal position.

3. *Though* is also more frequent in writing (with a weaker effect compared to *although*), when dialogic CCs are expressed, and if the subordinate clause is in final position.

4. *Even though* tends to be selected more often in the spoken mode, to express anticausal meaning, and when the subordinate clause is in sentence-final position.

5. Compared to L2 varieties, L1 varieties tend to use higher rates of *although* (at the expense of *though*), but it was argued that this pattern is mostly due to the idiosyncratic pattern found in IndE.

The fact that *although* responds somewhat more strongly to differences in mode of production lends some support to Quirk et al.'s (1985) claim that it is more formal than *though* (see also Biber et al. 1999, Huddleston & Pullum 2002, Aarts 1988). More fine-grained stylistic analyses would of course be required to substantiate this further. Associations between semantic types and particular conjunctions have thus far only been explored by Hilpert (2013a) and the author himself (Schützler 2017, 2018a). Conflicts between results in the present study and Hilpert's findings (e.g. concerning the connection between *although* and dialogic meaning) have been commented on before (e.g. in §10.3). On a methodological note, it is intuitively plausible that the strong link between the most frequent semantic type (dialogic) and the conjunction *although* can explain the high text frequency of this connective, as seen in Chapter 7. The case of *though* also supports this argument: If we ignore semantics (by controlling for this predictor), this conjunction appears to be of similar frequency as *even though*. If, however, we consider that *though* also associates strongly with the most frequent type (dialogic), we have the explanation for its rather high text frequency (again, see Chapter 7).

The conjunction *even though* stands out quite strongly from the other two, as it associates with the spoken mode and with the less frequent anticausal semantics. The semantic dimension explains why this marker has a much lower text frequency than *although* and *though* – again, this is not apparent from the analyses in Chapter 7. The association of *even though* with speech may be read as weakly supporting the claim that this conjunction has an emphatic character (suggested by the adverb *even*), which is sometimes made in the literature. In a vague sense, emphasis and more immediate modes of communication are characteristics of speech, rather than writing, but beyond this we can say little about the socio-stylistics of *even though*.

With regard to *clause structure*, i.e. the selection of a finite or nonfinite subordinate clause, neither the difference between L1 and L2 varieties nor the intra-constructional semantics of a CC seem to have systematic effects. The most important patterns are associated with mode of production, the positions of clauses and the subordinating conjunction, as follows:

1. Nonfinite subordinate clauses are more common in writing;

2. nonfinite clauses are more likely in nonfinal position; and

3. the conjunction *though* is most likely (and *even though* is least likely) to be followed by a nonfinite subordinate clause.

Nonfinite constructions are generally considered to be cognitively more complex since they imply, rather than overtly express, some of the necessary grammatical information. Their somewhat higher frequency in writing therefore comes as no surprise, since time constraints are considerably lower when producing and decoding written language.

The relationship between (non)finiteness and clause position is less straightforward. On the one hand, an association of nonfinite clauses with nonfinal positions does not seem ideal because it not only suspends the central (matrix-clause) proposition, but it additionally withholds grammatical information. Thus, AD/R has to store incomplete material at several levels until the gaps are filled in by the matrix clause. On the other hand, nonfinite structures are typically shorter than (and thus not as heavy as) finite structures. Quirk et al.'s (1985: 1036) notion of *resolution* would in this case predict that heavier (finite) structures should follow shorter (nonfinite) ones. In the present study, no hypotheses were attached to the possible correlation of clause positions and (non)finiteness. However, results suggest that a typical nonfinite CC presents the subordinate clause early and thus follows the principle of (sentence-level) end-weight, placing the matrix clause in focus position.

The strong link between the conjunction *though* and nonfinite subordinate clauses corresponds to findings by Hilpert (2013a), while findings concerning *even though* (which strongly favours finite clauses) and *although* (which is intermediate between the other two) are novel. The detected correlations result in a focusing of possible formal variants into more precisely defined subconstructions (see §12.3.3). It is the shortest marker (*though*) that is most likely to combine with nonfinite (and therefore, on average, also shorter) clauses, and the longest marker (*even though*) that is most likely to introduce finite (and thus longer) clauses. As a result, there will be a tendency to get short/nonfinite constructions

with *though*, long/finite constructions with *even though*, and intermediate constructions with *although*. While the effects of those correlations will in reality be quite subtle, it is noteworthy that, rather than balancing out the differences, the grammatical system seems to favour types of subconstructions that are more clearly differentiated in terms of length, or weight.

12.1.3 A marker-based summary

This brief section looks directly at each of the three conjunctions and highlights what contextual, functional and formal parameters they typically associate with. This is partly a semasiological view: Instead of asking what forms are typically selected, given a certain set of conditioning factors, it focuses on the typical functions, contexts of use, or concomitant formal characteristics of a given item – in this case a certain conjunction. This approach does not provide entirely novel insights but inverts the perspective on the results. In effect, it constitutes a brief return to the original point of view in Schützler (2018b), where the objective was to describe the differences between connectives, taking directly observable surface forms as the starting point for the analysis, without tying them into a more complex system of constructional choices, as in this volume.

The presentation in Figure 12.1 is based on information given in Figures 10.7–10.9 and 11.9–11.11, rescaling it in a standardised way.[3] Each level of the plot describes the associations of the three conjunctions with the two levels of a dichotomous variable across all nine varieties that were considered. If a conjunction is placed near the grey vertical line in the centre, it is relatively unresponsive to the respective variable. The further it is placed to the left or to the right, the greater its affinity to the condition indicated in the respective margin.

Concerning the association with L1 and L2 varieties of English, *although* tends towards the former, *though* tends towards the latter, and *even though* is very much indifferent to this dimension of variation. However, we saw (in all the relevant plots in §10.2) that IndE was exceptional in very strongly preferring the conjunction *though*. Since this highly erratic pattern did not correspond to any

---

[3]Take, for instance, the affinity of *even though* and anticausal CCs: Inspecting Figure 10.9, I determined the average rank of the top 36 slots and the average rank of the bottom 36 slots. The resulting values (18.5 and 54.5) were equated with 0 and 1, respectively, the ranks between them were rescaled accordingly, and the actual ranks of scenarios involving anticausal meanings were then placed on this standardised scale and plotted horizontally in Figure 12.1 for each conjunction. In concrete terms: If all 36 of the scenarios most favourable to the use of *even though* were anticausal in meaning, the symbol "E" would be placed on the very left of the plot. Conversely, if the 36 most favourable scenarios were all dialogic in meaning, the symbol would be placed on the very right. See the online appendix for details.

Figure 12.1: Association of markers with basic conditions; A = *although*,
T = *though*, E = *even though*

general tendency among L2 varieties, we must treat it (and its effect on the over-
all picture) with a certain suspicion. As discussed in the respective parts of the
analysis, the other results seem more reliable: *even though* typically occurs in
speech while the other two conjunctions associate with writing; *even though* is
much more likely if a CC has anticausal meaning, while both *although* and *though*
are much more common with dialogic CCs; *although* typically introduces subor-
dinate clauses in nonfinal position – the "grounding function" discussed by Al-
tenberg (1986: 22) – while clauses with the other two markers are more likely to
follow the matrix clause; and finally, *even though* is least likely and *though* most
likely to combine with nonfinite subordinate clauses, while *although* is interme-
diate in this regard. Evidently, the three markers pattern rather differently for
different factors – affinities between any two of them can be found for individual
variables but cannot be generalised. Differences between the three conjunctions
are complex and not easy to detect, since they require involved semantic and
syntactic analyses, and it is therefore unsurprising that the literature has thus
far lacked precise descriptions.

Based on these findings, the three conjunctions can nevertheless be shown in
their overall (dis)similarity. To this end, multi-dimensional scaling was applied
to the values plotted in Figure 12.1 above. Euclidean distances between the three
markers were calculated across the standardised values indicating their affini-
ties to different factors (see Footnote 3 on p. 209), and these values were then
reduced to coordinates in a two-dimensional space, as shown in Figure 12.2 –
see Schützler (2022) for the technical details involved in this procedure. The plot
shows three separate scenarios: One in which all five factors are included (va-
riety status, mode, semantics, clause position and subordinate clause structure;
see Figure 12.1 above), and two alternative scenarios in which one or two factors
are excluded, as indicated.

The general arrangement of conjunctions relative to each other is relatively
similar, irrespective of whether we base the analysis on the full set of variables or
on a subset: *although* and *though* are somewhat closer to each other, while *even*
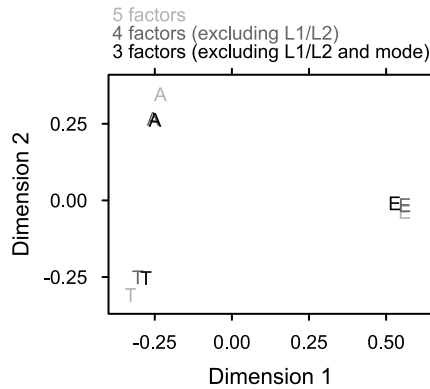
Figure 12.2: Similarities of conjunctions based on associations with predictors; A = *although*, T = *though*, E = *even though*

though stands apart. The respective distances for the three-factor scenario are 0.52 between *although* and *though*, 0.82 between *although* and *even though*, and 0.84 between *though* and *even though*, respectively.[4] The tentative claim made in Chapter 10 to the effect that *though* is more "multi-role" in character cannot be upheld from the perspective shown here: It may be true that *though* responds less sensitively to certain predictors than the other two conjunctions, but it is by no means positioned between them as regards general correlations with certain factor levels. The functional differentiation of the three conjunctions is of a much more complex and overlapping nature.

12.2 Wider contexts of investigation

This section outlines a few suggestions concerning potential directions for future work on CCs. Some of these are stock commentaries found at the end of any major book or research article, laying out what could have been done in an ideal world, with no restrictions on resources. Some of them are less promising and will accordingly be discussed rather briefly. Others, however, arise directly from the experience of this particular study and can be understood as serious suggestions for future work. Issues that concern the plausibility of the constructional choice model are reserved for §12.3 below.

Two possible expansions are briefly mentioned here, but not discussed at length, because, at least to the author, they seem ambitious beyond the manage-

---

[4]Note that these distances are based on the two-dimensional (potentially reductive) representation, not on the original underlying distances; again, see Schützler (2022).

able and branch out into domains far more general than English linguistics. The first concerns the discussion of concessives from a cross-linguistic perspective. Comparing the patterns that were found in the present study with patterns in other (Germanic) languages would be very much in the spirit of work by König (e.g. 1988, 1994, 2006), Kortmann (1996) and Rudolph (1996), for instance. This kind of undertaking would require historical and cross-linguistic expertise and would thus best be tackled collaboratively. Another aspect that branches out into far more general areas of knowledge concerns a more systematic investigation and categorisation of the topoi at work in anticausal and epistemic CCs (cf. §2.2.1 and §2.2.2). This could theoretically be undertaken not only for concessives but also for conditional and causal relations, since most of the implicational structures will be shared. Such an investigation would shed light on human cognition and the construction of a functional human world. However, knowledge of effects based on causes, results based on actions, or behaviours based on predispositions is psychologically so pervasive and basic, as well as culturally diverse, that it may well prove too vast an object of investigation. Similarly, the precise types of qualification and modification that operate between propositions in dialogic CCs (cf. §2.2.3) – which I called *themes* in this study – could be investigated more systematically. Like topoi, however, relations of this kind form an open class and establishing an inventory may well turn out to be a Sisyphean task.

12.2.1 From constructional subset to complete inventory

The present study was onomasiological in orientation, as semantic (and extralinguistic) functions were treated as primary and formal choices as secondary. However, since the analysis was restricted to CCs involving the three conjunctions *although*, *though* and *even though*, it remains unknown to what extent other means of encoding CCs are employed, for instance prepositional or coordinated constructions. Even more problematically, there are also constructions that cannot even be automatically retrieved from a corpus, since they generate concessive meaning purely from the content of propositions or from the discourse context. The problems involved in more comprehensive approaches to concessives are highlighted by Hoffmann (2005: 111; see also §1.3 above). However, casting the net wider in this way might be successful if the analysis was restricted to a corpus of suitable size – that is, a corpus large enough to contain a sufficiently wide range of constructions, but small enough for the researcher to essentially read it without recourse to automatic retrieval. Due to the enormous amount of manual work involved, this comprehensive approach would probably need to focus on a

limited range of varieties, and it would not generate enough material for meaningful register analyses. All of this, however, would depend on the resources that are invested.

Another problem that is faced when looking at all possible concessive constructions is that, unlike the complex sentences in the present study, they will in many cases not be syntactically equivalent. For instance, the clausal complements of conjunctions were classified as finite and nonfinite in this book, but this kind of classification does of course not apply to the complements of prepositions such as *despite* or *in spite of*. Other markers introduce their own specific complications, as, for instance, in the use of *notwithstanding* as a post- or preposition (Schützler 2018c); conjuncts like *however* or *nevertheless* would need to be treated differently; and "universal conditional-concessives" (e.g. *whatever you do*; *however hard we tried*; see §2.1.1) also evade the straightforward classifications applied in this study.

Finally, there is the question of syntactic structures that are perhaps not very frequent but quite salient, like the use of certain correlative markers, particularly in varieties beyond the Inner Circle (e.g. *although...but*; cf. §3.5). These appear as syntactic hybrids, since they combine coordinating and subordinating markers in a single construction. How exactly to classify those complex connectives – and whether to treat them as markers in their own right or as variants of existing subordinators – is very much an open question and has potential implications for syntactic theory.

Thus, there is certainly scope for expanding the focus of the present study and aiming at a fuller treatment of concessive constructions. The methodological challenges, however, are quite considerable, and the structure of a unified analytical framework would need to be developed along strongly modified or altogether different lines, compared to the present study.

12.2.2 Expanding the functional dimension

In a narrower sense, the function of a CC was defined at the interface of the two involved propositions, and it has been variously called a "semantic" and/or "pragmatic" function. The discussion of more complex views of CCs in this section looks beyond the construction but does not touch upon language-external (e.g. socio-stylistic) functions (but see §12.2.4 below).

The somewhat unsatisfactory results concerning clause position (see Chapter 9) raised the question of whether it is enough to look at the relation between propositions in a CC to predict formal realisations, or whether we should include the wider discourse context to this end. For instance, there was no systematic,

cross-varietal link between semantic types and the decision to place a subordinate clause in final or nonfinal position in the sentence. Particularly the hypothesis that the arrangement of clauses should be iconic of the semantic relation between propositions was not supported by the data. The general arrangement of component clauses might in fact be more systematically conditioned by content preceding or following the actual CC in question. For instance, does the proposition in one of the component clauses relate to propositions or arguments found earlier or later in the discourse, and is the respective clause therefore placed in proximity to those points of reference to increase textual cohesion and facilitate planning and processing? In order to address this issue we would in many cases have to look quite some distance to the left and right of a sentence to find clues that link the wider discourse to the respective CC. We would also need to categorise different kinds of anticipation in what precedes, different kinds of elaboration in what follows, their interactions, as well as instances in which no obvious discourse connection can be found. In other words: In addition to the intra-constructional relations identified in this study, we would need similar relations that apply to the wider discourse. Apart from being challenging at the coding stage, any such expansion would considerably increase the complexities of statistical models – that is, if a quantitative approach is still considered feasible under these circumstances in the first place. It was precisely for reasons like these that a discourse-analytic component was not included in the present study (see §1.1).

Even if we stay at the level of intra-constructional semantic/pragmatic functions as operationalised in this study, semantically ambiguous constructions (see §3.4) could be explicitly addressed in the analysis, for example via the addition of levels to the predictor variable TYPE (cf. §6.3.6). However, before this is considered, the epistemic type should be re-included, in spite of its relatively low overall frequency.

Thus, there is much that could be done concerning the expansion of the functional side of CCs. Like the expansion of candidate constructions discussed in §12.2.1, however, putting these ideas into practice would in many cases involve a considerable reworking of the analytic framework, particularly concerning quantitative methods.

12.2.3 The diachronic dimension

Section 2.1 provided the general historical background for the concessive class of adverbials in general and the conjunctions *although*, *though* and *even though* in particular, but the present study did not actively engage with the diachronic dimension of variation. The issues in filling this gap are once again mainly to do

with the availability of data and the amount of manual coding and disambiguation involved in the analysis. Diachronic work could make valuable contributions on several counts, as sketched in the following paragraphs.

One diachronic research question could concern the relatedness and development of semantic types of concessives. In particular, the more or less implicit treatment of anticausal CCs as somehow primary or prototypical and the associated notion that epistemic and dialogic CCs are derived from them (cf. Hilpert 2013a, Sweetser 1990) merit closer inspection. For instance, are dialogic CCs only notionally "later" than anticausal ones, in the sense that a "pragmaticalised" function is conceptualised as derived from a notionally more "logical" one? Or can we actually *show* that they appear later in the history of English? Further, if such diachronic processes can be traced: Do epistemic CCs take an intermediate position, or do they play some other role? Schützler's (2018a) study of AmE finds some evidence that all three conjunctions were more likely to carry dialogic meaning in the late 20th century, compared to the late 19th century. However, the statistical approach that was used is unlikely to stand the test of more rigorous methods, and only relatively weak tendencies were found. If confirmed, the diachronic derivation of epistemic meanings from anticausal meanings could be interpreted as a case of subjectification, because it results in a greater visibility of the active reasoning and inferencing of SP/W. This is remotely related to the well-known development of modal constructions from deontic to epistemic meanings, as discussed by Krug (2000: 91) in the context of grammaticalisation, for instance. The development of (putatively intersubjective) dialogic CCs would then be yet another step away from purely content-oriented readings. These possible diachronic trajectories would need to be investigated in new research efforts, however. Contra the notion of a diachronically increasing number of dialogic CCs, Burnham (1911: 33; cf. Footnote 17 on p. 21) suggests that it was quite common for Old English to use *þéah* – the predecessor of *though* – in a dialogic, quasi-adversative function (although Burnham does of course not use these terms). The secondary/derived status of the dialogic type can therefore not be taken for granted – on the contrary, it is not only possible that dialogic CCs have for a long time coexisted with anticausal (and epistemic) CCs, but they may actually have been the dominant (because more general) type to start with.

Another diachronic question concerns the changing fates of different markers concerning their availability. Frequency changes can shed light both on the grammaticalisation status of conjunctions and their stylistic values. Like the present study, such efforts would ideally use a complex framework that takes functional and formal parameters into account and thus goes beyond the mere measuring of text frequencies, as in Schützler (2018b: 165), for example.

12.2.4 More varieties and predictors?

The present study includes data from nine different varieties of English. On the whole, differences between those varieties were relatively slight or unsystematic. Including more varieties in follow-up studies would probably not shed more light on general differences between, say, inner-circle and outer-circle varieties, as those differences are apparently not particularly pronounced with regard to the phenomenon under investigation. Further, expanding the range of varieties would likely reveal more instances of idiosyncratic patterns that are hard to account for. Against this background, the inclusion of further (L2) varieties seems warranted only if there are specific, theoretically motivated expectations attached to those particular varieties. Such an expansion would then require a more careful consideration of the sociolinguistic realities in the respective territories. On the whole, however, CCs are probably not particularly salient and therefore play a relatively minor role as variety-based identity markers, as argued in §12.1 based on general patterns mostly characterised by inter-varietal similarity. However, the absence of US-American English from the investigated set of varieties constitutes a regrettable gap – as explained in §6.1, this is due to the incomplete status of ICE-USA. Rather than including more L2 varieties, a systematic comparison of BrE and AmE (i.e. English in the USA) might therefore be a valuable contribution. This would then need to be based on corpora beyond those from the ICE-family.

This study treated speech and writing as macro-stylistic categories and also considered what is involved in their production and processing. Getting a better idea of the stylistic value of different concessive markers would require a more fine-grained inspection of registers (or genres). ICE-corpora, however, are too small to investigate stylistic variation in detail, particularly if the construction is of medium frequency and is given a complex definition with multiple formal parameters, as in the present study. Distinguishing production-and-processing effects from truly stylistic effects will not always be easy and probably requires a careful operationalisation of genre. Truly social factors might also be of interest: If individual markers respond to differences in register – as claimed in some of the literature – they may also vary systematically between groups of speakers and writers. Once more, investigating this dimension of variation either requires corpora that include the appropriate metadata, or the adoption of more controlled (e.g. experimental) methodologies. On the whole, however, there seem to be several aspects of CCs more deserving of closer inspection than their precise sociolinguistic behaviour. Some have been outlined in this section as well as in §12.2.1 and §12.2.2 above, others will be discussed in the next section.

12.3 Concessives and Construction Grammar

In this study, CCs were conceptualised as a hierarchically organised system of choices. This section evaluates the success of the approach, suggests alternative approaches, and discusses theoretical implications.

12.3.1 Constructions as hierarchical choices

The hierarchical view of CCs in the choice model introduced in §4.1.3 means that we take a top-down perspective on constructions, with more general, broadly defined constructions at the top and more fully specified constructions at the bottom. This can be incorporated into an even more general hierarchy with three levels, the second and third of which comprise the choice model as implemented in this study: (i) language-external function, (ii) non-situational (or perhaps, language-internal) function, and (iii) form.

In the present study, language-external, socio-stylistic (or contextual) factors include the varieties themselves – perhaps grouped into L1 and L2 – and the two modes of production, writing and speech. Even if they are non-linguistic, factors like these can broadly be classified as functional: For instance, SP/W may consciously or unconsciously wish to flag up their association with a certain variety, and the selected formal realisations will then serve that function. Mode of production cannot be captured in exactly the same terms, as we can hardly claim that SP/W feels the need to express the fact that they are speaking (or writing). However, speaking and writing are clearly functions (or uses) that language can be put to, and they place certain constraints on how things are expressed. The consequent formal choices will thus again serve a higher function, namely making writing or speech work for both SP/W and AD/R. The relevant mechanisms can be viewed as production- or processing-related or as macro-stylistic (cf. §4.2). It is crucial to bear in mind that, as language-external factors, both variety and mode can potentially inform all lower-ranking properties of a construction.

One level below the two language-external parameters there are functions that cannot easily be linked to the situation or context in which language is produced. In the present study, only the semantic or pragmatic relationship between propositions within a CC was included at this level. Other functions of this kind, applicable to other constructions, could be found in the domain of modality (e.g. obligation meanings of different strengths), in other adverbial domains (e.g. temporal relations of different kinds) or in slight differences between semantic roles (e.g. different kinds of possessor-possessum relationships). Such functions have in common that, while they do of course refer to some real-world situation or

some relation external to the linguistic form, there is no immediate link with the situation (or the conditions) under which language is produced. An indicator for the identification of such intermediate functions may be the question of whether they are rooted in the socio-stylistic or cognitive characteristics of the situation or directly linked to what SP/W wishes to express. In the present study, for example, constructing an anticausal CC is not motivated socio-stylistically but from a specific message that needs to be conveyed.

Chapter 8 explored the notion that the number of CCs of different semantic types may vary across modes of production and different varieties of English. However, the general view taken of CCs in Chapters 9, 10 & 11 was that the construction proper only begins at the intermediate, message-related level of the semantic function, which then finds expression in the various possible formal realisations. The construction is thus implicitly treated as context-free: If we look at a CC in isolation, we can identify its internal semantic make-up, the arrangement of clauses, the conjunction that is involved, as well as the structure of the subordinate clause. Delimiting a CC in this way results in what was called a *hermetic* view (see §4.1.3): All parameters relevant for the analysis of the construction can be recovered from its propositional content and its form.

The formal level is then comprised of (i) clause position, (ii) a marker, and (iii) the syntactic structure of the subordinate clauses. These properties not only rank lowest in the general hierarchy (EXTRALINGUISTIC FUNCTION → SEMANTIC FUNCTION → FORM), but they can also be ranked internally. In this study, I took the view that the primary, highest-level decision concerns where to place the component clauses, which are, after all, the largest elements involved. Next, the link between the two clauses (i.e. the conjunction) was given precedence over the formal realisation of the subordinate clause, which was motivated from traditional grammatical thinking whereby the clause depends upon (or complements) its conjunction.

If we accept the notion that form follows function, we can still question the assumption that certain formal properties are conditioned by others in a unidirectional (or hierarchical) way. In other words: Can we even identify formal properties of different ranks, or should we treat form as a single, multi-dimensional component of a construction? In the context of the present study, for instance, is it reasonable to frame the dependency as POSITION → MARKER → CLAUSAL STRUCTURE? Or would a partly different arrangement, or the rejection of any hierarchy, be more plausible? These issues will be discussed in the following section.

12.3.2 Alternative models of constructional variation

A critique of the choice model of constructional variation (cf. §4.1.3) can be based on a particular view of usage-based CxG, as outlined in §4.1.2 and illustrated in Figure 4.1. The schema introduced there involves the three formal properties of CCs along with the semantic/pragmatic dimension. These parameters are all put on an equal footing, as shown by the lines that establish all possible cross-connections. In this model, it does not seem contradictory to assume that it is the semantic function that drives formal variation and the establishment of typical formal patterns: Meaning is primary and needs to be formally expressed, and we therefore have an unavoidable function-to-form hierarchy (see §12.3.1 above). However, a usage-based assumption would be that certain formal correlations (e.g. more instances of *although* in nonfinal position; cf. Chapter 10) become cognitively strengthened simply through their co-occurrence. That is, instead of being guided by some mechanism that works its way through the different formal layers in a top-down fashion, SP/W intuitively accesses all relevant formal levels at the same time, producing a formally complex construction whose internal dependencies (apart from FUNCTION → FORM) are not even theoretically relevant. The contrast between the two views is shown in Figure 12.3. In the hierarchical model in panel (a), fully language-external factors like VARIETY and MODE as well as semantic or pragmatic factors have an impact on all formal aspects of a CC. Concerning formal parameters, however, the model implies that SP/W has stored inventories of likely realisations at different levels of granularity, from a general syntactic grid of sequentially ordered component clauses via the selection of a subordinator to the eventual realisation of the subordinate clause as finite or nonfinite. These could then be called *subconstructions*, with more specific ones nested in more general (or schematic) ones. In panel (b), on the other hand, external and semantic/pragmatic factors have a direct impact on a single (if still multi-faceted) formal choice.

As a model of real-time language production, the model in panel (a) of Figure 12.3 seems less efficient, as it suggests that SP/W accesses the different layers of a construction in a sequential way, going through a chain of decisions. Perspective (b), on the other hand, is more economical and therefore plausible, since all formal properties are directly accessed in bulk. However, even in this holistic view the internal relation between formal properties still needs to be established – we still want to take a look into the black box that contains POSITION, MARKER and CLAUSE in panel (b) in order to find out how its content is patterned.

To resolve the conflict between the two panels in Figure 12.3, I will argue that they simply make two different contributions to answering the same question,
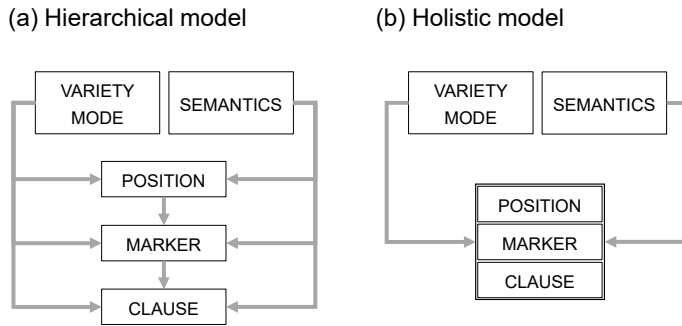
Figure 12.3: Hierarchical (a) vs holistic (b) views of formal dependencies in CCs

namely: What are the typically expected formal properties of a CC, given a particular semantic function and context of production? The two components of Figure 12.3 approach this issue from two perspectives. Panel (a) represents a particular view of how subconstructions are organised at the formal level, proceeding from higher-level, general syntactic grids to the more local properties of sentence-internal linkage and the structure of embedded clauses. This schematic and idealised view is directly aligned with the quantitative analyses in this book, based on regression models that become increasingly complex as we move from the most general and schematic subconstructions to more fully specified ones. Panel (b), on the other hand, establishes a more direct link between functions (both extra- and intra-linguistic) and forms and treats the latter as more unitary. This reflects that SP/W does of course make a single choice when encoding a CC in a particular situation.

The three equations shown in (88) return to the syntax of statistical models that were used in Chapters 9, 10 & 11, in order to further illustrate the hierarchical thinking that was applied. Note that several variables are given more general names here (e.g. SEMANTICS, POSITION and CLAUSE), and that random parts are not restated, because they are irrelevant for the discussion at hand. The logic of these related models is that the higher-level outcomes POSITION and MARKER become predictor variables at lower levels. In a sense, constructional variation (involving three formal parameters) is operationalised as three separate alternations, with a single formal parameter as the outcome in each case.

(88)   Statistical models in the hierarchical perspective

```
position ~ mode * semantics
marker   ~ mode * (semantics + position)
clause   ~ mode * (semantics + position + marker)
```

Above, it was argued that the two components of Figure 12.3 merely take two perspectives on essentially the same question, and that the hierarchical view breaks the holistic view up into more manageable units (i.e. binary or ternary alternations) but otherwise serves the same purpose. The three formulations in (88) show that this is not strictly true. For instance, POSITION can impact upon MARKER, and MARKER can impact upon CLAUSE, but not vice versa. Based on this design and on results shown in Chapter 11, we can argue that using the conjunction *though* makes a nonfinite subordinate clause more likely, but it does not follow that using a nonfinite subordinate clause makes using the conjunction *though* more likely. An ideal model, however, should perhaps accommodate both views: Neither does SP/W select a certain type of clause on the basis of a certain marker, nor is the marker selected on the basis of a clause type, but the two of them are selected together. Not only can we question the exact hierarchy of formal levels, but we can question the very idea of hierarchies. What, then, would be the methodological consequences of a truly holistic perspective concerning the formal side of CCs? The three equations in (89) show options that will be discussed in more detail below.

(89)   Possible models for the holistic perspective

```
form                       ~ mode * semantics
form_func                  ~ mode
position | marker | clause ~ mode * semantics
```

The outcome variable FORM in the first model has an exceptional structure: Its levels correspond to all twelve possible discrete combinations of clause positions (×2), conjunctions (×3) and clause structures (×2). Using only extra-linguistic and semantic predictors, probabilities of these outcome variants could theoretically be predicted using a multinomial model, but this is indeed no more than a theoretical possibility: Ternary outcomes are already difficult enough to handle (see supplementary materials for Chapter 10; see also Fahy et al. 2022), and analyses with more than three outcome categories (as in Schützler & Herzky 2021) are very much the exception. However, based on such models we could easily focus on subconstructions at a more general level, for instance by comparing the cumulated proportion of all formal variants that involve the marker *though* to the cumulated proportion of formal variants involving the other two conjunctions, or by comparing the cumulated proportion of all variants with subordinate clauses in final position to the respective value for clauses in nonfinal position.[5]

---

[5]Note that formal parameters subsumed into the outcome variable – here, as well as in the count models discussed below – have to be categorical; continuous characteristics (like clause length) must either be excluded or converted into categories.

To a similar effect, a count model (cf. Chapters 7 & 8) could be employed to measure the rates of occurrence of variants, to be then converted into proportions or percentages – see, for instance, the approach in Schützler (2022). The respective (theoretical) outcome variable is labelled FORM_FUNC in the second line of (89): We need to include the functional (i.e. semantic) dimension in the outcome that is counted, because only properties of text units can be used as predictors. This approach is computationally easier to handle than the complex multinomial model sketched above, but it should in principle yield the same results. This, too, comes with its own complications, however. For instance, counts need to be established for each outcome category in each text. Given the number of $n = 4{,}902$ texts in the present study, we would need $n = 24$ observations per text (12 forms × 2 semantics), blowing the dataset up to $n = 117{,}648$ observations.[6] In analogy to the approach described for the multinomial model, typical subconstructions can be captured by summing up the estimated rates for specific outcomes across contrasting broader categories (e.g. final vs nonfinal; *although* vs *though* vs *even though*).

A third option is the application of regression modelling with a multivariate outcome (e.g. Afifi et al. 2020, Johnson & Wichern 2007), as indicated in the third equation in (89): Separate regressions are formulated for each formal parameter (in this case: POSITION, MARKER and CLAUSE), probably using identical predictor structures for theoretical reasons. However, the model not only contains information about the relations between outcomes and predictors, but also knows about the correlations between levels of the three outcomes.

Finally, an approach that does not strictly fall into the domain of regression is Structural Equation Modelling (e.g. Hoyle 2012, Kaplan 2009). Among other things, the flexible and powerful procedures that it provides can account for correlated dependents and complex interrelationships between all variables involved. However, a fuller discussion of these techniques cannot be provided here.

The alternative modelling strategies sketched in this section have in common that they target subconstructions not in a hierarchical way – as shown schematically in (88) – but holistically. The relative benefits of such techniques remain to be tested. Their discussion, however, highlights that the present study has proposed merely one particular perspective on constructional variation, which must not be taken as the final word but as a point of departure for future approaches to constructional variation (and, perhaps, change). In particular, such approaches

---

[6]Compare the count models in Chapters 7 & 8, which had "only" $n = 14{,}706$ and $n = 19{,}608$ observations, respectively, as described in Appendices B.1 & B.2 and documented in the published data (Schützler 2021).

would reconcile the view of constructions as formally holistic with the fact that they can also break down into syntactically more schematic (or, less fully specified) subconstructions. Notionally, the holistic view of constructions was already a driving force behind the analyses in the present volume. Quantitatively, however, it could be implemented more rigorously in future research.

12.3.3 Processes in constructional variation and change

Beyond fundamental quantitative issues involved in the analysis of formally and functionally complex constructions, future research should take a closer look at the relation between function and form. What, for instance, defines a subconstruction? What specific processes can be involved in what we popularly refer to as *constructionalisation*? And, again, how can we support our analysis of such processes using our quantitative toolkit? This section makes some suggestions concerning these points. The focus is on the constructions that were the topic of this volume (CCs), but the notions that are developed have wider applicability.

A basic, common-sense assumption is that the relationship between functional and formal variation is not random but structured. There is at the very least a tendency for certain (e.g. semantic) functions to be expressed using certain (e.g. syntactic) forms. This kind of correlation can be extended to the relations between different formal parameters as well, as explained earlier. While these notions are in fact relatively theory-neutral and do not in themselves advance Construction Grammar, their theoretical relevance is strengthened if we link them to production and processing. Both are assumedly facilitated if a certain function correlates with distinct formal properties; conversely, they are made more difficult if the relation between function and form is fuzzy and unsystematic. In other words: If there is a lot of overlap between the formal means used to express different functions, decoding a message will be cognitively more challenging. In the study at hand, the simplified functional view focused on anticausal and dialogic meanings. In this context, constructionalisation would consist of a tidier mapping of particular formal properties onto each of the two semantics. Measuring the discreteness of form-function patterns would then be one concern of quantitative Construction Grammar.

The concept that is introduced below will be called *constructional focusing*. It was also used in Schützler (2018b: 124–128), alongside two other concepts, "standardised constructional difference" and "constructional specialisation". As compared to the present study, the original framing of constructional focusing is somewhat problematic: On the one hand, the terms "focusing" and "specialisation" are relatively similar; on the other hand, the earlier study was semasio-

logical (form-driven), not onomasiological (function-driven), in outlook. I would suggest that it is more efficient to use a single concept developed strictly from the perspective of functions.

Figure 12.4 draws a schematic comparison between relatively unfocused (or "diffuse") and relatively focused constructions. It suggests that function and form are continuous dimensions. While this is theoretically true, functional and formal *categories* are more likely to be used in practice, as in the present volume.
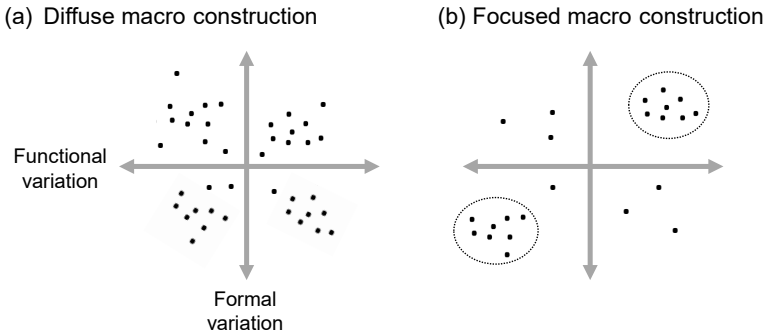
Figure 12.4: Diffuse and focused macro constructions

In Figure 12.4a, instances of a previously defined macro construction (e.g. CCs) are scattered rather randomly across the available functional and formal space. As semantic properties vary (horizontally), formal properties also vary (vertically), but not in a particularly systematic way: The same formal means are available when encoding different functions, and the function-to-form mapping is therefore relatively *diffuse*. In Figure 12.4b, on the other hand, variation is much more structured: Functions on the left are expressed by formal means not available to functions on the right to the same extent. A large number of form-function pairings are still theoretically possible (and will therefore appear in data), but the probabilities are high for certain combinations and low for others, and formal overlap between different functions is much more limited. We would therefore speak of a macro construction that has focused into relatively distinct subconstructions.

This concept of focusing nicely dovetails with the notion that constructions (and thus, subconstructions) are certain combinations of form and function that are processed, stored and produced by language users to generate meaning. If, in an exemplar-based representation, certain parameters at different levels combine more often than others, they are strengthened and will be cognitively more readily accessed, as discussed in §4.1.2. This would apply both to constructions

with fairly general, productive grammatical properties (e.g. CCs) and more idiosyncratic and unpredictable constructions like the covariational-conditional construction discussed in §4.1.1, for instance.

In quantitative terms, measuring constructional focusing would require the kind of holistic approach outlined in §12.3.2. It would perhaps be most interesting to trace diachronic changes in focusing and thus the emergence of subconstructions, but variation may well exist between national varieties of English or genres, too. In cognitive terms, changing (or systematically varying) degrees of constructional focusing would shed light on how language users classify and store mental (e.g. semantic/pragmatic) categories, and how they relate them to linguistic forms.

12.4 In conclusion

The constructions under investigation in this book turned out to be relatively stable across varieties as regards the general constraints that regulate them; speech and writing, on the other hand, have a greater impact on formal variation. On the whole, however, CCs seem to be most interesting if we look at them with a focus on the coherence and interplay of intra-constructional functional and formal parameters. Semantics, clause positions, the specific conjunctions themselves as well as the syntactic realisations of subordinate clauses – all of these interact in a systematic way and tend to form a structured set of subconstructions within the broader class of CCs. The precise relationships between functional and formal facets in these and other (possibly rather different) constructions deserve even closer attention in the future, and research efforts of this kind can contribute to the development of new theories and quantitative methods in Construction Grammar. The present volume has provided a few pointers in this direction.

Appendix A: Data summary (ICE)

A.1 Corpus design

Traditional ICE components consist of approximately 1 million words (60% spoken, 40% written) in 500 texts of approx. 2,000 words each, sampled according to a standardised scheme as described in §6.1 (see also https://www.ice-corpora. uzh.ch/en.html).

The first column in Table A.1 states the alphanumerical handle that is conventionally used to identify macro genres in ICE, along with an explanation in parentheses. The section (spoken or written) that a text belongs to is indicated by the first capital letter ("S" or "W") in those handles. Subsections, like sections, are not indicated as an extra level in Table A.1 but become clear from the structure: Macro genres S1A and S1B together make up the spoken subsection of dialogues (S1); similarly, macro genres S2A and S2B constitute the subsection of monologues (S2); macro genres W1A and W1B comprise non-printed writing (W1); and macro genres W2A–F constitute printed writing (W2). The number of texts conventionally sampled for each genre within the macro genres is given in the second column. Note that in the published components of ICE, texts are counted consecutively within macro genres, so that "ICE-GB:S1B-017" is a classroom lesson, "ICE-GB:S1B-024" is a broadcast discussion, and "ICE-GB:S1B-049" is a broadcast interview, for instance. Further note that ICE-Nigeria contains considerably more than 500 texts, because the compilers did not follow the practice of combining several shorter texts of the same kind into larger units of 2,000 word. However, ICE-Nigeria still adheres to the total number of 1 million words.

The specific genre labels are given in the third column of Table A.1, and the final column states their abbreviations as used to code the data for the variable GENRE for the statistical analyses in the present study (cf. §6.3.6). These particular abbreviations were introduced for ICE-Nigeria (and were also adopted by the compilers of ICE-Scotland), because they are more explicit than the purely alphanumerical labels traditionally used in ICE.

Table A.1: The structure of the *International Corpus of English*

Macro genre	*n* (texts)	Genre	Short label
S1A (private dialogues)	90	face-to-face conversations	con
	10	phonecalls	ph
S1B (public dialogues)	20	classroom lessons	les
	20	broadcast discussions	bdis
	10	broadcast interviews	bint
	10	parliamentary debates	parl
	10	legal cross-examinations	cr
	10	business transactions	btrans
S2A (unscripted monologues)	20	spontaneous commentaries	com
	30	unscripted speeches	unsp
	10	demonstrations	dem
	10	legal presentations	leg
S2B (scripted monologues)	20	broadcast news (20)	bnew
	20	broadcast talks (20)	btal
	10	non-broadcast talks (10)	nbtal
W1A (student writing)	10	student essays (10)	ess
	10	exam scripts (10)	ex
W1B (letters)	15	social letters (15)	sl
	15	business letters (15)	bl
W2A (academic writing)	10	humanities (10)	AHum
	10	social sciences (10)	ASoc
	10	natural sciences (10)	ANat
	10	technology (10)	ATec
W2B (popular writing)	10	humanities (10)	PHum
	10	social sciences (10)	PSoc
	10	natural sciences (10)	PNat
	10	technology (10)	PTec
W2C (reportage)	20	press news reports (20)	rep
W2D (instructional writing)	10	administrative writing (10)	adm
	10	skills/hobbies (10)	SkHo
W2E (persuasive writing)	10	press editorials (10)	ed
W2F (creative writing)	20	novels & short stories (20)	nov

A.2 Word counts

Word counts (and corpus queries) are based on files in which content captured by the following tags was ignored: <X>…</X>, <0>…</0>, <&>…</&> (see comments in §6.1). For ICE-NIG, word counts as stated in the corpus manual are reported.[1]

Table A.2: Word counts in components of ICE

	Number of words		
Variety	spoken	written	Total
BrE	647,837	432,769	1,080,606
IrE	627,627	423,757	1,051,384
CanE	612,060	393,099	1,005,159
AusE	655,117	423,868	1,078,985
JamE	614,860	403,665	1,018,525
NigE	609,586	400,796	1,010,382
IndE	674,678	412,523	1,087,201
HKE	709,365	471,734	1,181,099
SingE	600,635	402,506	1,003,141
Total	5,751,765	3,764,717	9,516,482

A.3 Token numbers

Table A.3 reports the number of tokens (per conjunction) that were semantically disambiguated, using the original four semantic types (anticausal, dialogic, epistemic and narrow-scope dialogic), and not counting tokens that were excluded. Frequencies are reported for both speech and writing. The total number of tokens is $n = 3,502$. This forms the basis of Models A & B in Chapters 7 & 8 – see the files "concessives_1" and "concessives_2" in the published data (Schützler 2021).

Table A.4 shows the reduced number of tokens used for the remaining analyses (Models C–E in Chapters 9–11), in which epistemic and narrow-scope concessives were excluded. The remaining total number of tokens was $n = 3,275$ – see the file "concessives_3" in the published data (Schützler 2021).

Finally, Table A.5 documents the precise number of tokens, sorted by the categories relevant in the most complex scenario analysed in Chapter 11: (i) variety, (ii) mode of production, (iii) semantic type, (iv) subordinate clause position, (v) conjunction and (vi) clause structure. The total number of tokens is $n = 3,275$.

---

[1]Available at https://sourceforge.net/projects/ice-nigeria/; accessed 14 October 2023.

Table A.3: Token numbers of conjunctions in ICE (Models A & B); W = written; S = spoken

Variety	although			though			even though		
	W	S	Total	W	S	Total	W	S	Total
GB	219	131	350	118	61	179	29	16	45
IRE	140	64	204	57	22	79	30	32	62
CAN	166	59	225	54	17	71	33	35	68
AUS	172	70	242	65	20	85	23	33	56
JAM	93	74	167	60	25	85	24	35	59
NIG	73	40	113	56	51	107	26	46	72
IND	69	29	98	129	106	235	25	26	51
SING	131	74	205	82	30	112	27	25	52
HK	212	77	289	81	18	99	51	41	92
Total	1,275	619	1,893	702	350	1,052	268	289	557

Table A.4: Token numbers of conjunctions in ICE (Models C–E); W = written; S = spoken

Variety	although			though			even though		
	W	S	Total	W	S	Total	W	S	Total
GB	210	123	333	102	56	158	28	14	42
IRE	136	62	198	49	21	70	30	31	61
CAN	162	57	219	47	13	60	31	35	66
AUS	168	67	235	56	17	73	21	32	53
JAM	88	71	159	49	19	68	24	29	53
NIG	69	36	105	50	49	99	26	44	70
IND	65	27	92	110	95	205	24	26	50
SING	125	69	194	73	27	100	27	23	50
HK	202	74	276	78	18	96	50	40	90
Total	1,225	586	1,811	614	315	929	261	274	535

Table A.5: Detailed list of tokens by all categories; S = spoken, W = written, a = anticausal, d = dialogic, fn = final, nf = nonfinal, A = *although*, T = *though*, E = *even though*, f = finite, n = nonfinite

BrE / IrE / CanE

			BrE S a	BrE S d	BrE W a	BrE W d	IrE S a	IrE S d	IrE W a	IrE W d	CanE S a	CanE S d	CanE W a	CanE W d
fn	A	f	2	58	11	9	3	28	7	52	3	36	6	45
		n	1	3		2			1	3				
	T	f	1	39	5	4		15	1	33	1	8	5	19
		n		3						1				2
	E	f	6	3	13	6	14	9	9	10	17	6	14	8
		n						1						
nf	A	f	12	45	88	96	7	18	7	56	3	15	18	82
		n		2		4	1	5	3	7			2	9
	T	f	3	9	56	27	1	3	2	9		3	3	12
		n		1	2	8		2		3	1	1		5
	E	f	2	3	7	2	3	4	3	8	4	8	3	6
		n												

AusE / JamE / NigE

			AusE S a	AusE S d	AusE W a	AusE W d	JamE S a	JamE S d	JamE W a	JamE W d	NigE S a	NigE S d	NigE W a	NigE W d
fn	A	f	3	38	4	80	3	27	7	16		17		15
		n		1		1		1						
	T	f	1	10	4	28		10	1	14	1	25	1	16
		n			1	3		1				3		1
	E	f	10	8	9	6	7	5	6	7	7	15	9	11
		n							1					1
nf	A	f	3	19	14	66	6	32	9	47	5	14	4	48
		n	2	1		3	1	1	3	6				2
	T	f	3	3	2	13	1	6	2	19	3	15	4	24
		n			3	2		1	4	9		2	1	3
	E	f	5	8	3	3	9	8	2	8	3	19	1	4
		n		1										

IndE / SingE / HKE

			IndE S a	IndE S d	IndE W a	IndE W d	SingE S a	SingE S d	SingE W a	SingE W d	HKE S a	HKE S d	HKE W a	HKE W d
fn	A	f	4	9	3	14	2	21	6	32	4	19	7	53
		n										1		1
	T	f	14	35	10	32	1	9	3	26	2	6	9	22
		n		3		3		1		1		2		3
	E	f	6		5	7	8	2	12	8	13	8	21	12
		n												
nf	A	f	5	9	8	35	6	36	12	69	9	41	25	111
		n			3	2	1	3		6			1	4
	T	f	15	27	13	39	1	10	7	26	2	5	15	21
		n	1		4	9	3	2	1	9	1		5	3
	E	f	9	11	5	6	3	10	1	6	5	14	7	10
		n			1									

Appendix B: Statistical models

B.1 Model A: Frequencies of conjunctions (Chapter 7)

The analysis is based on a total number of n = 14,706 individual observations, which represent the frequency values of the three markers (*although*, *though* and *even though*) in all n = 4,902 texts contained in the nine corpora – see the file "concessives_1" in the published data (Schützler 2021). Table B.1 shows the overall token numbers as well as the number of levels of the random factor GENRE. The model was run with four chains, each with n = 3,500 iterations and a warmup phase of n = 1,000 iterations. The number of data points in the posterior sample was thus n = 10,000. The R-hat diagnostic equalled 1.00 for all parameters, indicating the convergence of the four chains. The full model output (i.e. tables of coefficients) can be found in the online materials (cf. §1.4). Priors are shown in Table B.2. The prior for the intercept was fixed at the geometric mean of the frequencies of the most frequent item in ICE-GB (*the*, with a normalised frequency of $f \approx$ 64,000 pmw) and an assumed rare (lexical) item with f = 1 pmw. The standard deviation of the prior was then set to the value of 1.8, so that the difference between the mean and the two extreme values mentioned above equalled roughly three standard deviations. As seen in Table B.1, there are no missing genres (nor, of course, texts) in any of the nine varieties, since zero counts and positive counts were all entered into the underlying data frame.

B.2 Model B: Frequencies of semantic types (Chapter 8)

This analysis is based on a total number of n = 19,608 individual observations, which represent the frequencies of all four semantic types (anticausal, epistemic, wide-scope dialogic and narrow-scope dialogic) in the n = 4,902 texts that make up the nine components of ICE used in this study. The complete data are accessible in the file "concessives_2" in the published data (Schützler 2021). All nine regression models were run with four chains, each with n = 3,500 iterations and a warmup phase of n = 1,000 iterations. The number of data points in the posterior sample was therefore n = 10,000. The R-hat diagnostic equalled 1.00 for

Table B.1: Number of observations for all models; *n* = total number of observations, GENRE/TEXT = number of levels of the respective random factor

| | Model | | | | | |
| | A | | B | | C–E | | |
Variety	*n*	GENRE	*n*	GENRE	*n*	GENRE	TEXT
GB	1,500	32	2,000	32	533	32	262
IRE	1,500	32	2,000	32	329	31	189
CAN	1,500	32	2,000	32	345	32	207
AUS	1,500	32	2,000	32	361	32	194
JAM	1,500	32	2,000	32	280	30	177
NIG	2,706	32	3,608	32	274	31	189
IND	1,500	32	2,000	32	347	31	207
SING	1,500	32	2,000	32	344	31	190
HK	1,500	32	2,000	32	462	31	251

all parameters, which indicates that the four chains reached convergence. The respective parts of Tables B.1 & B.2 document the total number of observations, the number of levels of the random factor GENRE, as well as the priors that were specified. More comprehensive documentation in the form of regression tables can be found online (cf. §1.4). The selection of priors followed the same principles as for Model A (cf. Appendix B.1 above). Like for Model A, there are no missing genres (or missing texts) in any of the nine varieties, since zero counts were entered into the data frame along with positive counts.

B.3 Model C: Clause position (Chapter 9)

These analyses are based on a total of *n* = 3,275 individual observations, distributed across the nine varieties under investigation. The data can be found in the file "concessives_3" in what has been published as (Schützler 2021). The total number of observations as well as the number of levels of the two random factors GENRE and TEXT are given in Table B.1, the priors are documented in Table B.2, and full regression tables for all individual models can be retrieved from the online repository (cf. §1.4). A separate model with four chains of *n* = 4,750 iterations each and a warmup phase of *n* = 1,000 iterations was run for each variety. The resulting number of posterior samples was thus *n* = 15,000 per model.

The R-hat diagnostic equalled 1.00 for all parameters, which is an indicator of the convergence of the four chains. There are missing genres in six out of the nine varieties, i.e. genres which do not appear as a level of the cluster variable GENRE, because they produced no hits: IrE (missing: business transactions), JamE (administrative writing and cross examinations), NigE (parliamentary debates), IndE (cross examinations), SingE and HKE (phone calls).

B.4 Model D: Choice of conjunction (Chapter 10)

The analysis is based on a total number of n = 3,275 individual observations made in all nine varieties under investigation. The data are contained in the file "concessives_3" in Schützler (2021). This model was run with four chains, each with n = 5,500 iterations and a warmup phase of n = 1,000 iterations, which resulted in n = 18,000 posterior samples. The R-hat diagnostic took the value of 1.00 for all parameters, confirming the convergence of chains. Basic information on token numbers and priors are provided in Tables B.1 & B.2, and – as for the other models – much more detailed model summaries are documented online (cf. §1.4). Missing genres (and missing texts) are the same as in Model C (see Appendix B.3 above).

B.5 Model E: Clause structure (Chapter 11)

Like Models C & D, this analysis is based on n = 3,275 individual observations in the nine varieties that were included. The data can be accessed in the file "concessives_3" in the published data (Schützler 2021). All of the nine models were run with four chains, each with n = 4,000 iterations and a warmup of n = 1,000 iterations, which yielded a total of n = 12,000 posterior samples. R-hat equalled 1.00 for all parameters. Again, see Tables B.1 & B.2 for basic information on overall token numbers and the number of levels of the two random factors GENRE and TEXT, as well as for information regarding the priors. Full documentation of model output is provided in the online appendix (cf. §1.4). Once again, missing genres (and missing texts) are the same as in Models C & D (see Appendices B.3 & B.4 above).

Table B.2: Specification of priors for all models

Model	Parameter	Prior
A	Intercept	normal (−8.28, 1.8)
	b	normal (0, 1.5)
	SD	normal (0, 1.5)
	cor	lkj(2)
B	Intercept	normal (−8.28, 1.8)
	b	normal (0, 1.5)
	SD	normal (0, 1.5)
	cor	lkj(2)
C	Intercept	normal (0, 1)
	b	normal (0, 1.5)
	SD	normal (0, 1)
	cor	lkj(2)
D	Intercept (*though*)	normal (−0.5, 2)
	Intercept (*even though*)	normal (−0.5, 2)
	b	normal (0, 1.5)
	SD (*though*)	normal (0, 1.5)
	SD (*even though*)	normal (0, 1.5)
	cor	lkj(2)
E	Intercept	normal (−3, 2)
	b	normal (0, 1.5)
	SD	normal (0, 1)
	cor	lkj(2)

References

Aarts, Bas. 1988. Clauses of concession in written present-day British English. *Journal of English Linguistics* 21(1). 39–58. DOI: 10.1177/007542428802100104.

Aarts, Bas. 2011. *Oxford Modern English Grammar.* Oxford: Oxford University Press.

Aarts, Bas, Joanne Close, Geoffrey Leech & Sean Wallis (eds.). 2013. *The verb phrase in English: Investigating recent language change with corpora* (Studies in English Language). Cambridge: Cambridge University Press. DOI: 10.1017/CBO9781139060998.

Abercrombie, David. 1991. RP today: Its position and prospects. In David Abercrombie (ed.), *Fifty years in phonetics: Selected papers*, 48–53. Edinburgh: Edinburgh University Press.

Afifi, Abdelmonem, Susanne May, Robin A. Donatello & Virginia A. Clark. 2020. *Practical multivariate analysis* (Texts in Statistical Science). Boca Raton: Chapman & Hall/CRC. DOI: 10.1201/9781315203737.

Agresti, Alan. 2013. *Categorical data analysis* (Wiley Series in Probability and Statistics). 3rd edn. Hoboken: Wiley.

Akinnaso, F. Niyi. 1982. On the differences between spoken and written language. *Language and Speech* 25(2). 97–125. DOI: 10.1177/002383098202500201.

Altenberg, Bengt. 1986. Contrastive linking in spoken and written English. In Gunnel Tottie & Ingegerd Bäcklund (eds.), *English in speech and writing: A symposium* (Studia Anglistica Upsaliensia 60), 13–40. Stockholm: Almqvist & Wiksell.

Anscombre, Jean-Claude. 1989. Théorie de l'argumentation, topoï, et structuration discursive. *Revue Québécoise de Linguistique* 18(1). 13–55. DOI: 10.7202/602639ar.

Anthony, Laurence. 2018. *AntConc: A freeware corpus analysis toolkit for concordancing and text analysis.* Version 3.5.7. http://www.laurenceanthony.net/software.html (7 December, 2021).

Azar, Moshe. 1997. Concessive relations as argumentations. *Text & Talk* 17(3). 301–316. DOI: 10.1515/text.1.1997.17.3.301.

References

Baayen, R. Harald. 2008. *Analyzing linguistic data: A practical introduction to statistics using R.* Cambridge: Cambridge University Press. DOI: 10.1017/CBO9780511801686.

Baker, Paul. 2009. The BE06 corpus of British English and recent language change. *International Journal of Corpus Linguistics* 14(3). 312–337. DOI: 10.1075/ijcl.14.3.02bak.

Baker, Paul. 2010–2011. *The AmE2006 reference corpus.* Lancaster: Lancaster University.

Barth, Dagmar. 2000. "that's true, although not really, but still": Expressing concession in spoken English. In Elizabeth Couper-Kuhlen & Bernd Kortmann (eds.), *Cause – condition – concession – contrast: Cognitive and discourse perspectives* (Topics in English Linguistics 33), 411–437. Berlin: Mouton de Gruyter. DOI: 10.1515/9783110219043.4.411.

Barth-Weingarten, Dagmar. 2003. *Concession in spoken English: On the realisation of a discourse-pragmatic relation* (Language in Performance 28). Tübingen: Narr.

Bates, Douglas, Martin Maechler, Ben Bolker & Steven Walker. 2020. *lme4: Linear mixed-effects models using Eigen and S4.* R package version 1.1–23. https://cran.r-project.org/web/packages/lme4/lme4.pdf (7 December, 2021).

Benveniste, Émile. 1971. *Problems in general linguistics* (Miami Linguistics Series 8). Coral Gables: University of Miami Press.

Bergs, Alexander & Gabriele Diewald. 2008. Introduction: Constructions and language change. In Alexander Bergs & Gabriele Diewald (eds.), *Constructions and language change* (Trends in Linguistics. Studies and Monographs 194), 1–21. Berlin: Mouton de Gruyter. DOI: 10.1515/9783110211757.1.

Berlage, Eva. 2009. Prepositions and postpositions. In Günter Rohdenburg & Julia Schlüter (eds.), *One language, two grammars? Differences between British and American English* (Studies in English Language), 130–148. Cambridge: Cambridge University Press. DOI: 10.1017/CBO9780511551970.007.

Berlage, Eva. 2014. *Noun phrase complexity in English* (Studies in English Language). Cambridge: Cambridge University Press. DOI: 10.1017/CBO9781139057684.

Bernaisch, Tobias. 2015. *The lexis and lexicogrammar of Sri Lankan English* (Varieties of English Around the World G54). Amsterdam: John Benjamins. DOI: 10.1075/veaw.g54.

Best, Henning & Christof Wolf. 2015. Logistic regression. In Henning Best & Christof Wolf (eds.), *The Sage Handbook of Regression Analysis and Causal Inference*, 153–172. London: Sage. DOI: 10.4135/9781446288146.

Biber, Douglas. 1988. *Variation across speech and writing*. Cambridge: Cambridge University Press. DOI: 10.1017/CBO9780511621024.

Biber, Douglas & Susan Conrad. 2009. *Register, genre, and style* (Cambridge Textbooks in Linguistics). Cambridge: Cambridge University Press. DOI: 10.1017/9781108686136.

Biber, Douglas, Stig Johansson, Geoffrey Leech, Susan Conrad & Edward Finegan. 1999. *Longman Grammar of Spoken and Written English*. London: Longman.

Bolstad, William M. & James M. Curran. 2017. *Introduction to Bayesian statistics*. 3rd edn. Hoboken: Wiley. DOI: 10.1002/9781118593165.

Borkin, Ann. 1980. On some conjuncts signalling dissonance in written expository English. *Studia Anglica Posnaniensia* 12. 47–59.

Brinton, Laurel J. & Donna M. Brinton. 2010. *The linguistic structure of Modern English*. Amsterdam: John Benjamins. DOI: 10.1075/z.156.

Bryant, Margaret M. (ed.). 1962. *Current American usage*. New York: Funk & Wagnalls.

Bürkner, Paul-Christian. 2020. *brms: Bayesian regression models using Stan*. R package version 2.16.1. https://cran.r-project.org/web/packages/brms/brms.pdf (4 September, 2021).

Burnham, Josephine M. 1911. *Concessive constructions in Old English prose* (Yale Studies in English 39). New York: Henry Holt.

Buschfeld, Sarah, Thomas Hoffmann, Magnus Huber & Alexander Kautzsch (eds.). 2014. *The evolution of Englishes: The Dynamic Model and beyond* (Varieties of English Around the World G49). Amsterdam: John Benjamins. DOI: 10.1075/veaw.g49.

Buschfeld, Sarah & Alexander Kautzsch. 2017. Towards an integrated approach to postcolonial and non-postcolonial Englishes. *World Englishes* 36(1). 104–126. DOI: 10.1111/weng.12203.

Bybee, Joan L. 2001. *Phonology and language use* (Cambridge Studies in Linguistics 94). Cambridge: Cambridge University Press. DOI: 10.1017/CBO9780511612886.

Bybee, Joan L. 2006. From usage to grammar: The mind's response to repetition. *Language* 82(4). 711–33.

Bybee, Joan L. 2010. *Language, usage and cognition*. Cambridge: Cambridge University Press. DOI: 10.1017/CBO9780511750526.

Bybee, Joan L. 2013. Usage-based theory and exemplar representations of constructions. In Thomas Hoffmann & Graeme Trousdale (eds.), *The Oxford Handbook of Construction Grammar*, 49–69. Oxford: Oxford University Press. DOI: 10.1093/oxfordhb/9780195396683.013.0004.

References

Cameron, A. Colin & Pravin K. Trivedi. 2013. *Regression analysis of count data* (Econometric Society Monographs 53). Cambridge: Cambridge University Press. DOI: 10.1017/CBO9781139013567.

Cappelle, Bert. 2006. Particle placement and the case for "allostructions". *Constructions* Special volume 1. 1–28. DOI: 10.24338/cons-381.

Chafe, Wallace L. 1976. Givenness, contrastiveness, definiteness, subjects, topics and point of view. In Charles N. Li (ed.), *Subject and topic*, 27–55. New York: Academic Press.

Chafe, Wallace L. 1982. Integration and involvement in speaking, writing, and oral literature. In Deborah Tannen (ed.), *Spoken and written language: Exploring orality and literacy* (Advances in Discourse Processes 9), 35–53. Norwood: Ablex.

Chafe, Wallace L. 1984. How people use adverbial clauses. *Proceedings of the Annual Meeting of the Berkeley Linguistics Society* 10. 437–449. DOI: 10.3765/bls.v10i0.1936.

Chafe, Wallace L. 1985. Linguistic differences produced by differences between speaking and writing. In David R. Olson, Nancy Torrance & Angela Hildyard (eds.), *Literacy, language and learning: The nature and consequences of reading and writing*, 105–123. Cambridge: Cambridge University Press.

Chafe, Wallace L. 1994. *Discourse, consciousness, and time: The flow and displacement of conscious experience in speaking and writing*. Chicago: University of Chicago Press.

Chafe, Wallace L., John W. Du Bois & Sandra A. Thompson. 1991. Towards a new corpus of spoken American English. In Karin Aijmer & Bengt Altenberg (eds.), *English corpus linguistics*, 64–82. London: Longman.

Chafe, Wallace L. & Jane Danielewicz. 1987. Properties of spoken and written language. In Rosalind Horowitz & S. Jay Samuels (eds.), *Comprehending oral and written language*, 83–113. San Diego: Academic Press.

Chen, Guohua. 2000. The grammaticalization of concessive markers in Early Modern English. In Olga Fischer, Annette Rosenbach & Dieter Stein (eds.), *Pathways of change: Grammaticalization in English* (Studies in Language Companion Series 53), 87–110. Amsterdam: John Benjamins. DOI: 10.1075/slcs.53.06che.

Clyne, Michael. 1992. Pluricentric languages: Introduction. In Michael Clyne (ed.), *Pluricentric languages: Differing norms in different nations* (Contributions to the Sociology of Language 62), 1–9. Berlin: Mouton de Gruyter. DOI: 10.1515/9783110888140.1.

Couper-Kuhlen, Elizabeth & Bernd Kortmann (eds.). 2000a. *Cause – condition – concession – contrast: Cognitive and discourse perspectives* (Topics in English Linguistics 33). Berlin: Mouton de Gruyter. DOI: 10.1515/9783110219043.

Couper-Kuhlen, Elizabeth & Bernd Kortmann. 2000b. Introduction. In Elizabeth Couper-Kuhlen & Bernd Kortmann (eds.), *Cause – condition – concession – contrast: Cognitive and discourse perspectives* (Topics in English Linguistics 33), 1–8. Berlin: Mouton de Gruyter. DOI: 10.1515/9783110219043-001.

Couper-Kuhlen, Elizabeth & Sandra A. Thompson. 2000. Concessive patterns in conversation. In Elizabeth Couper-Kuhlen & Bernd Kortmann (eds.), *Cause – condition – concession – contrast: Cognitive and discourse perspectives* (Topics in English Linguistics 33), 381–410. Berlin: Mouton de Gruyter. DOI: 10.1515/9783110219043-016.

Crawford, William J. 2009. The mandative subjunctive. In Günter Rohdenburg & Julia Schlüter (eds.), *One language, two grammars? Differences between British and American English* (Studies in English Language), 257–276. Cambridge: Cambridge University Press. DOI: 10.1017/CBO9780511551970.015.

Crevels, Mily. 2000. Concessives on different semantic levels: A typological perspective. In Elizabeth Couper-Kuhlen & Bernd Kortmann (eds.), *Cause – condition – concession – contrast: Cognitive and discourse perspectives* (Topics in English Linguistics 33), 313–339. Berlin: Mouton de Gruyter. DOI: 10.1515/9783110219043-013.

Croft, William. 2005. Logical and typological arguments for Radical Construction Grammar. In Jan-Ola Östman & Mirjam Fried (eds.), *Construction grammars: Cognitive grounding and theoretical extensions* (Constructional Approaches to Language 3), 273–314. Amsterdam: John Benjamins. DOI: 10.1075/cal.3.11cro.

Croft, William & D. Alan Cruse. 2004. *Cognitive linguistics* (Cambridge Textbooks in Linguistics). Cambridge: Cambridge University Press. DOI: 10.1017/CBO9780511803864.

Culicover, Peter W. 1976. *Syntax*. New York: Academic Press.

Cumming, Geoff & Robert Calin-Jageman. 2017. *Introduction to the new statistics: Estimation, open science, and beyond.* New York: Routledge. DOI: 10.4324/9781315708607.

Davies, Mark. 2010. *The Corpus of Historical American English: 400 million words, 1810–2009.* https://www.english-corpora.org/coha/ (7 December, 2021).

Di Meola, Claudio. 1998. Zur Definition einer logisch-semantischen Kategorie: Konzessivität als „versteckte Kausalität". *Linguistische Berichte* 175. 329–352.

Di Meola, Claudio. 2004. Ikonische Beziehungen zwischen Konzessivrelation und Konzessivkonnektoren. In Hardarik Blühdorn, Eva Breindl & Ulrich Hermann Waßner (eds.), *Brücken schlagen: Grundlagen der Konnektorensemantik* (Lin-

References

guistik: Impulse & Tendenzen 5), 287–308. Berlin: Mouton de Gruyter. DOI: 10.1515/9783110890266.287.

Diessel, Holger. 2005. Competing motivations for the ordering of main and adverbial clauses. *Linguistics* 43(3). 449–470. DOI: 10.1515/ling.2005.43.3.449.

Diessel, Holger. 2008. Iconicity of sequence: A corpus-based analysis of the positioning of temporal adverbial clauses in English. *Cognitive Linguistics* 19(3). 457–482. DOI: 10.1515/COGL.2008.018.

Du Bois, John W., Wallace L. Chafe, Charles Meyer, Sandra A. Thompson, Robert Englebretson & Nii Martey. 2000–2005. *Santa Barbara Corpus of Spoken American English, Parts 1–4*. Philadelphia: Linguistic Data Consortium.

Edwards, Alison. 2016. *English in the Netherlands: Functions, forms and attitudes* (Varieties of English Around the World G56). Amsterdam: John Benjamins. DOI: 10.1075/veaw.g56.

Eitle, Hermann. 1914. *Die Satzverknüpfung bei Chaucer* (Anglistische Forschungen 44). Heidelberg: Carl Winter.

Fahy, Matthew, Jesse Egbert, Benedikt Szmrecsanyi & Douglas Biber. 2022. Comparing logistic regression, multinomial regression, classification trees and random forests applied to ternary variables: Three-way genitive variation in English. In Ole Schützler & Julia Schlüter (eds.), *Data and methods in corpus linguistics: Comparative approaches*, 194–223. Cambridge: Cambridge University Press. DOI: 10.1017/9781108589314.008.

Ferguson, Charles A. 1982. Foreword. In Braj B. Kachru (ed.), *The other tongue: English across cultures*, vii–xi. Urbana: University of Illinois Press.

Ferraresi, Gisella (ed.). 2011. *Konnektoren im Deutschen und im Sprachvergleich* (Studien zur deutschen Sprache 53). Tübingen: Narr.

Fillmore, Charles J. 1988. The mechanisms of "Construction Grammar". *Proceedings of the Annual Meeting of the Berkeley Linguistics Society* 14. 35–55. DOI: 10.3765/bls.v14i0.1794.

Fowler, Roger. 1991. *Language in the news: Discourse and ideology in the press*. London: Routledge.

Francis, W. Nelson & Henry Kučera. 1979. *Manual of information to accompany A Standard Corpus of Present-Day Edited American English, for use with digital computers*. 3rd edn. Providence: Department of Linguistics, Brown University. http://korpus.uib.no/icame/manuals/brown (7 December, 2021).

Fried, Mirjam. 2008. Constructions and constructs: Mapping a shift between predication and attribution. In Alexander Bergs & Gabriele Diewald (eds.), *Constructions and language change* (Trends in Linguistics. Studies and Monographs 194), 47–79. Berlin: Mouton de Gruyter. DOI: 10.1515/9783110211757.47.

Gelman, Andrew & Sander Greenland. 2019. Are confidence intervals better termed "uncertainty intervals"? *British Medical Journal* 366(15381). 1–2. DOI: 10.1136/bmj.l5381.

Gelman, Andrew & Jennifer Hill. 2007. *Data analysis using regression and multi-level/hierarchical models* (Analytical Methods for Social Research). Cambridge: Cambridge University Press. DOI: 10.1017/CBO9780511790942.

Givón, Talmy. 1990. *Syntax: A functional-typological introduction*, vol. 2. Amsterdam: John Benjamins. DOI: 10.1075/z.50.

Goldberg, Adele E. 2003. Constructions: A new theoretical approach to language. *Trends in Cognitive Sciences* 7(5). 219–224. DOI: 10.1016/S1364-6613(03)00080-9.

Goldberg, Adele E. 2006. *Constructions at work: The nature of generalization in language.* Oxford: Oxford University Press. DOI: 10.1093/acprof:oso/9780199268511.001.0001.

Goldberg, Adele E. 2011. Corpus evidence of the viability of statistical preemption. *Cognitive Linguistics* 22(1). 131–153. DOI: 10.1515/cogl.2011.006.

Goldberg, Adele E. 2013. Constructionist approaches. In Thomas Hoffmann & Graeme Trousdale (eds.), *The Oxford Handbook of Construction Grammar*, 15–31. Oxford: Oxford University Press. DOI: 10.1093/oxfordhb/9780195396683.013.0002.

Görlach, Manfred. 1990. *Studies in the history of the English language* (Anglistische Forschungen 210). Heidelberg: Carl Winter.

Granger, Sylviane. 1996. Learner English around the world. In Sidney Greenbaum (ed.), *Comparing English worldwide: The International Corpus of English*, 13–24. Oxford: Clarendon Press.

Granger, Sylviane, Estelle Dagneaux, Fanny Meunier & Magali Paquot. 2009. *International Corpus of Learner English: Version 2.* Louvain-la-Neuve: Presses universitaires de Louvain.

Grant, William. 1914. *The pronunciation of English in Scotland* (Cambridge Primers of Pronunciation). 2nd edn. Cambridge: Cambridge University Press.

Greenbaum, Sidney. 1988. A proposal for an international computerized corpus of English. *World Englishes* 7(3). 315. DOI: 10.1111/j.1467-971X.1988.tb00241.x.

Greenbaum, Sidney. 1991. ICE: The International Corpus of English. *English Today* 7(4). 3–7. DOI: 10.1017/S0266078400005836.

Greenbaum, Sidney. 1992. A new corpus of English: ICE. In Jan Svartvik (ed.), *Directions in corpus linguistics* (Trends in Linguistics. Studies and Monographs 65), 171–179. Berlin: Mouton de Gruyter. DOI: 10.1515/9783110867275.171.

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

References

Greenbaum, Sidney & Jan Svartvik. 1990. The London-Lund Corpus of Spoken English. In Jan Svartvik (ed.), *The London-Lund Corpus of Spoken English: Description and research* (Lund Studies in English 82), 11–59. Lund: Lund University Press.

Halliday, Michael A. K. & Ruqaiya Hasan. 1976. *Cohesion in English* (English Language Series 9). London: Longman.

Halliday, Michael A. K. & Christian M. I. M. Matthiessen. 2004. *An introduction to functional grammar*. 3rd edn. London: Arnold. DOI: 10.4324/9780203783771.

Hancil, Sylvie. 2014. The final particle *but* in British English: An instance of cooptation and grammaticalization at work. In Sylvie Hancil & Ekkehard König (eds.), *Grammaticalization: Theory and data* (Studies in Language Companion Series 162), 235–255. Amsterdam: John Benjamins. DOI: 10.1075/slcs.162.12han.

Harris, Martin. 1988. Concessive clauses in English and Romance. In John Haiman & Sandra A. Thompson (eds.), *Clause combining in grammar and discourse* (Typological Studies in Language 18), 71–99. Amsterdam: John Benjamins. DOI: 10.1075/tsl.18.06har.

Hawkins, John A. 1990. A parsing theory of word order universals. *Linguistic Inquiry* 21(2). 223–261.

Hawkins, John A. 1992. Syntactic weight versus information structure in word order variation. *Linguistische Berichte* (Sonderhefte, Band 4): *Informationsstruktur und Grammatik*. 196–219.

Hawkins, John A. 1994. *A performance theory of order and constituency* (Cambridge Studies in Linguistics 73). Cambridge: Cambridge University Press. DOI: 10.1017/CBO9780511554285.

Hawkins, John A. 2000. The relative order of prepositional phrases in English: Going beyond manner – place – time. *Language Variation and Change* 11(3). 231–266. DOI: 10.1017/S0954394599113012.

Heine, Bernd & Tanya Kuteva. 2002. *World lexicon of grammaticalization*. Cambridge: Cambridge University Press.

Hermodsson, Lars. 1978. *Semantische Strukturen der Satzgefüge im kausalen und konditionalen Bereich* (Studia Anglistica Upsaliensia 18). Uppsala: Almqvist & Wiksell.

Hermodsson, Lars. 1994. Der Begriff „konzessiv": Terminologie und Analysen. *Studia Neophilologica* 66(1). 59–75. DOI: 10.1080/00393279408588131.

Hilpert, Martin. 2007. Just because it's new doesn't mean people will notice it. *English Today* 23(3–4). 29–33. DOI: 10.1017/S0266078407003070.

Hilpert, Martin. 2013a. *Constructional change in English: Developments in allomorphy, word formation, and syntax* (Studies in English Language). Cambridge: Cambridge University Press. DOI: 10.1017/CBO9781139004206.

Hilpert, Martin. 2013b. Corpus-based approaches to constructional change. In Thomas Hoffmann & Graeme Trousdale (eds.), *The Oxford Handbook of Construction Grammar*, 458–475. Oxford: Oxford University Press. DOI: 10.1093/oxfordhb/9780195396683.013.0025.

Hoffmann, Sebastian. 2005. *Grammaticalization and English complex prepositions: A corpus-based study* (Routledge Advances in Corpus Linguistics). London: Routledge. DOI: 10.4324/9780203008225.

Hoffmann, Thomas & Lucia Siebers (eds.). 2009. *World Englishes: Problems, properties and prospects (Selected papers from the 13th IAWE conference)* (Varieties of English Around the World G40). Amsterdam: John Benjamins. DOI: 10.1075/veaw.g40.

Hoffmann, Thomas & Graeme Trousdale (eds.). 2013. *The Oxford Handbook of Construction Grammar*. Oxford: Oxford University Press. DOI: 10.1093/oxfordhb/9780195396683.001.0001.

Hopper, Paul J. 1991. On some principles of grammaticalization. In Elizabeth C. Traugott & Bernd Heine (eds.), *Approaches to grammaticalization*, vol. 1 (Typological Studies in Language 19:1), 17–36. Amsterdam: John Benjamins. DOI: 10.1075/tsl.19.1.04hop.

Hopper, Paul J. & Elizabeth C. Traugott. 2003. *Grammaticalization* (Cambridge Textbooks in Linguistics). 2nd edn. Cambridge: Cambridge University Press. DOI: 10.1017/CBO9781139165525.

Hox, Joop. 2010. *Multilevel analysis: Techniques and applications* (Quantitative Methodology Series). 2nd edn. London: Routledge. DOI: 10.4324/9780203852279.

Hoyle, Rick H. (ed.). 2012. *Handbook of structural equation modeling.* New York: Guildford Press.

Huddleston, Rodney D. & Geoffrey K. Pullum. 2002. *The Cambridge Grammar of the English Language.* Cambridge: Cambridge University Press. DOI: 10.1017/9781316423530.

Hundt, Marianne. 2004–2013. *The B-Brown-1931 Corpus.* Zürich: University of Zürich.

Hundt, Marianne & Ulrike Gut (eds.). 2012. *Mapping unity and diversity worldwide: Corpus-based studies of new Englishes* (Varieties of English Around the World G43). Amsterdam: John Benjamins. DOI: 10.1075/veaw.g43.

Hundt, Marianne, Andrea Sand & Paul Skandera. 1999. *Manual of information to accompany the Freiburg-Brown Corpus of American English ("Frown").* Freiburg: Albert-Ludwigs-Universität Freiburg, Englisches Seminar. http://korpus.uib.no/icame/manuals/frown (7 December, 2021).

References

Izutsu, Mitzuko N. & Katsunobu Izutsu. 2014. "Final hanging *but*" in American English: Where a formal coordinator meets a functional subordinator. In Sylvie Hancil & Ekkehard König (eds.), *Grammaticalization: Theory and data* (Studies in Language Companion Series 162), 257–285. Amsterdam: John Benjamins. DOI: 10.1075/slcs.162.13izu.

Jenkins, Jennifer. 2000. *The phonology of English as an international language: New models, new norms, new goals.* Oxford: Oxford University Press.

Jenkins, Jennifer. 2007. *English as a lingua franca: Attitude and identity.* Oxford: Oxford University Press.

Jenkins, Jennifer. 2015. *Global Englishes: A resource book for students* (Routledge English Language Introductions). London: Routledge. DOI: 10.4324/9781315761596.

Johansson, Stig, Geoffrey N. Leech & Helen Goodluck. 1978. *Manual of information to accompany the Lancaster-Oslo/Bergen Corpus of British English, for use with digital computers.* University of Oslo, Department of English. http://korpus.uib.no/icame/manuals/lob.

Johnson, Daniel E. 2014. Progress in regression: Why natural language data calls for mixed-effects models. https://web.archive.org/web/20220808180352/http://www.danielezrajohnson.com/johnson_2014b.pdf (7 December, 2021).

Johnson, Richard A. & Dean W. Wichern. 2007. *Applied multivariate statistical analysis.* 6th edn. Upper Saddle River: Pearson Prentice-Hall.

Jucker, Andreas (ed.). 1993. *The noun phrase in English: Its structure and variability* (Anglistik & Englischunterricht 49). Heidelberg: C. Winter.

Kachru, Braj B. 1985. Standards, codification and sociolinguistic realism: The English language in the outer circle. In Randolph Quirk & Henry G. Widdowson (eds.), *English in the world: Teaching and learning the language and literatures*, 11–30. Cambridge: Cambridge University Press.

Kachru, Braj B. 1988. The sacred cows of English. *English Today* 4(4). 3–8. DOI: 10.1017/S0266078400000973.

Kaplan, David. 2009. *Structural equation modeling: Foundations and extensions* (Advanced Quantitative Techniques in the Social Sciences 10). 2nd edn. Los Angeles: Sage. DOI: 10.4135/9781452226576.

Kirk, John & Gerald Nelson. 2018. The International Corpus of English project: A progress report. *World Englishes* 37(4). 697–716. DOI: 10.1111/weng.12350.

Kjellmer, Göran. 2009. The revived subjunctive. In Günter Rohdenburg & Julia Schlüter (eds.), *One language, two grammars? Differences between British and American English* (Studies in English Language), 246–256. Cambridge: Cambridge University Press. DOI: 10.1017/CBO9780511551970.014.

Koch, Peter & Wulf Oesterreicher. 1985. Sprache der Nähe – Sprache der Distanz: Mündlichkeit und Schriftlichkeit im Spannungsfeld von Sprachtheorie und Sprachgeschichte. *Romanistisches Jahrbuch* 36(1). 15–43. DOI: 10.1515/9783110244922.15.

König, Ekkehard. 1985. On the history of concessive connectives in English: Diachronic and synchronic evidence. *Lingua* 66(1). 1–19. DOI: 10.1016/S0024-3841(85)90240-2.

König, Ekkehard. 1988. Concessive connectives and concessive sentences: Cross-linguistic regularities and pragmatic principles. In John A. Hawkins (ed.), *Explaining language universals*, 145–166. Oxford: Blackwell.

König, Ekkehard. 1991a. Concessive relations as the dual of causal relations. In Dietmar Zaefferer (ed.), *Semantic universals and universal semantics* (Groningen-Amsterdam Studies in Semantics 12), 190–209. Berlin: Foris Publications. DOI: 10.1515/9783110870527-010.

König, Ekkehard. 1991b. Konzessive Konjunktionen. In Arnim von Stechow & Dieter Wunderlich (eds.), *Semantics: An international handbook of contemporary research* (Handbooks of Linguistics and Communication Science 6), 631–639. Berlin: Mouton de Gruyter. DOI: 10.1515/9783110126969.7.631.

König, Ekkehard. 1994. Concessive clauses. In R. E. Asher (ed.), *The Encyclopedia of Language and Linguistics*, vol. 2, 679–681. Oxford: Pergamon Press.

König, Ekkehard. 2006. Concessive clauses. In Keith Brown (ed.), *Encyclopedia of Language and Linguistics*, vol. 2, 820–824. Amsterdam: Elsevier.

König, Ekkehard & Peter Eisenberg. 1984. Zur Pragmatik von Konzessivsätzen. In Gerhard Stickel (ed.), *Pragmatik in der Grammatik: Jahrbuch 1983 des Instituts für Deutsche Sprache* (Sprache der Gegenwart 60), 313–332. Düsseldorf: Schwann-Bagel.

König, Ekkehard & Peter Siemund. 2000. Causal and concessive clauses: Formal and semantic relations. In Elizabeth Couper-Kuhlen & Bernd Kortmann (eds.), *Cause – condition – concession – contrast: Cognitive and discourse perspectives* (Topics in English Linguistics 33), 341–360. Berlin: Mouton de Gruyter. DOI: 10.1515/9783110219043.4.341.

König, Ekkehard & Elizabeth C. Traugott. 1988. Pragmatic strengthening and semantic change: The conventionalizing of conversational implicature. In Werner Hüllen & Rainer Schulze (eds.), *Understanding the lexicon: Meaning, sense and world knowledge in lexical semantics* (Linguistische Arbeiten 210), 110–124. Tübingen: Niemeyer. DOI: 10.1515/9783111355191.110.

Kortmann, Bernd. 1996. *Adverbial subordination: A typology and history of adverbial subordinators based on European languages* (Empirical Approaches to Language Typology 18). Berlin: Mouton de Gruyter. DOI: 10.1515/9783110812428.

References

Kortmann, Bernd, Kerstin Lunkenheimer & Katharina Ehret (eds.). 2020. *The Electronic World Atlas of Varieties of English*. DOI: 10.5281/zenodo.3712132. https://ewave-atlas.org (7 December, 2021).

Kortmann, Bernd & Benedikt Szmrecsanyi. 2011. Parameters of morphosyntactic variation in World Englishes: Prospects and limitations of searching for universals. In Peter Siemund (ed.), *Linguistic universals and language variation* (Trends in Linguistics. Studies and Monographs 231), 264–290. Berlin: Mouton de Gruyter. DOI: 10.1515/9783110238068.264.

Krifka, Manfred. 2008. Basic notions of information structure. *Acta Linguistica Hungarica* 55(3–4). 243–276. DOI: 10.1556/ALing.55.2008.3-4.2.

Krug, Manfred. 2000. *Emerging English modals: A corpus-based study of grammaticalization* (Topics in English Linguistics 32). Berlin: Mouton de Gruyter. DOI: 10.1515/9783110820980.

Kruschke, John K. 2015. *Doing Bayesian data analysis: A tutorial with R, JAGS, and Stan*. 2nd edn. Amsterdam: Elsevier.

Laitinen, Mikko. 2010. Describing "orderly differentiation": Compiling the corpus of English in Finland. *English Today* 26(1). 26–33. DOI: 10.1017/S0266078409990459.

Lambrecht, Knud. 1996. *Information structure and sentence form: Topic, focus, and the mental representations of discourse referents* (Cambridge Studies in Linguistics 71). Cambridge: Cambridge University Press. DOI: 10.1017/CBO9780511620607.

Lang, Ewald. 2000. Adversative connectors on distinct levels of discourse: A re-examination of Eve Sweetser's three-level approach. In Elizabeth Couper-Kuhlen & Bernd Kortmann (eds.), *Cause – condition – concession – contrast: Cognitive and discourse perspectives* (Topics in English Linguistics 33), 235–256. Berlin: Mouton de Gruyter. DOI: 10.1515/9783110219043-010.

Langacker, Ronald W. 1985. Observations and speculations on subjectivity. In John Haiman (ed.), *Iconicity in syntax* (Typological Studies in Language 6), 109–150. Amsterdam: John Benjamins. DOI: 10.1075/tsl.6.07lan.

Langacker, Ronald W. 1987. *Foundations of cognitive grammar*, vol. 1: *Theoretical prerequisites*. Stanford: Stanford University Press.

Langacker, Ronald W. 1988. A usage-based model. In Brygida Rudzka-Ostyn (ed.), *Topics in cognitive linguistics* (Current Issues in Linguistic Theory 50), 127–161. Amsterdam: John Benjamins. DOI: 10.1075/cilt.50.06lan.

Langacker, Ronald W. 1990. Subjectification. *Cognitive Linguistics* 1(1). 5–38. DOI: 10.1515/cogl.1990.1.1.5.

Leech, Geoffrey & Nicholas Smith. 2005. Extending the possibilities of corpus-based research in the twentieth century: A prequel to LOB and FLOB. *ICAME Journal* 29. 83–98.

Leitner, Gerhard. 1992. English as a pluricentric language. In Michael Clyne (ed.), *Pluricentric languages: Differing norms in different nations* (Contributions to the Sociology of Language 62), 179–237. Berlin: Mouton de Gruyter. DOI: 10.1515/9783110888140.179.

Lenker, Ursula. 2010. *Argument and rhetoric: Adverbial connectors in the history of English* (Topics in English Linguistics 64). Berlin: Mouton de Gruyter. DOI: 10.1515/9783110216066.

Linell, Per. 2005. *The written language bias in linguistics: Its nature, origins and transformations* (Routledge Advances in Communication and Linguistic Theory). London: Routledge. DOI: 10.4324/9780203342763.

Long, J. Scott. 1997. *Regression models for categorical and limited dependent variables* (Advanced Quantitative Techniques in the Social Sciences 7). Thousand Oaks: Sage.

Long, J. Scott. 2015. Regression models for nominal and ordinal outcomes. In Henning Best & Christof Wolf (eds.), *The Sage Handbook of Regression Analysis and Causal Inference*, 173–204. London: Sage. DOI: 10.4135/9781446288146.

Luke, Douglas A. 2004. *Multilevel modeling* (Quantitative Applications in the Social Sciences 143). Thousand Oaks: Sage. DOI: 10.4135/9781412985147.

May, Peter. 2012. *The Lewis Man.* London: Quercus.

McArthur, Tom. 1979. The status of English in and furth of Scotland. In Adam J. Aitken & Tom McArthur (eds.), *Languages of Scotland* (Occasional Papers of the Association for Scottish Literary Studies 4), 50–67. Edinburgh: Chambers.

McArthur, Tom. 1987. The English Languages? *English Today* 3(3). 9–11. DOI: 10.1017/S0266078400013511.

McArthur, Tom. 1998. *The English languages.* Cambridge: Cambridge University Press. DOI: 10.1017/9780511621048.

McArthur, Tom. 2003. World English, Euro-English, Nordic English. *English Today* 19(1). 54–58. DOI: 10.1017/S026607840300107X.

McElreath, Richard. 2020. *Statistical rethinking: A Bayesian course with examples in R and Stan* (Texts in Statistical Science). 2nd edn. Boca Raton: CRC Press. DOI: 10.1201/9780429029608.

Mesthrie, Rajend. 2003. The World Englishes paradigm and contact linguistics: Refurbishing the foundations. *World Englishes* 22(4). 449–461. DOI: 10.1111/j.1467-971X.2003.00312.x.

References

Mesthrie, Rajend & Rakesh M. Bhatt. 2008. *World Englishes: The study of new linguistic varieties* (Key Topics in Sociolinguistics). Cambridge: Cambridge University Press. DOI: 10.1017/CBO9780511791321.

Meurman-Solin, Anneli. 2001. Genre as a variable in socio-historical linguistics. *European Journal of English Studies* 5(2). 241–256. DOI: 10.1076/ejes.5.2.241.7311.

Miller, Jim & Regina Weinert. 1998. *Spontaneous spoken language: Syntax and discourse.* Oxford: Clarendon Press.

Modiano, Marko. 1999. Standard English(es) and educational practices for the world's lingua franca. *English Today* 15(4). 3–13. DOI: 10.1017/S0266078400011196.

Moessner, Lilo. 2001. Genre, text type, style, register: A terminological maze? *European Journal of English Studies* 5(2). 131–138. DOI: 10.1076/ejes.5.2.131.7312.

Molenberghs, Geert & Geert Verbeke. 2005. *Models for discrete longitudinal data* (Springer Series in Statistics). New York: Springer. DOI: 10.1007/0-387-28980-1.

Mondorf, Britta. 2004. *Gender differences in English syntax* (Linguistische Arbeiten 491). Tübingen: Niemeyer. DOI: 10.1515/9783110943375.

Mukherjee, Joybrato, Marco Schilk & Tobias Bernaisch. 2010. Compiling the Sri Lankan component of ICE: Principles, problems, prospects. *ICAME Journal* 34. 64–77.

Mulder, Jean & Sandra A. Thompson. 2008. The grammaticization of *but* as a final particle in English conversation. In Laury Ritva (ed.), *Crosslinguistic studies of clause combining: The multifunctionality of conjunctions* (Typological Studies in Language 80), 179–204. Amsterdam: John Benjamins. DOI: 10.1075/tsl.80.09mul.

Mulder, Jean, Sandra A. Thompson & Cara Penry Williams. 2009. Final *but* in Australian English conversation. In Pam Peters, Peter Collins & Adam Smith (eds.), *Comparative studies in Australian and New Zealand English: Grammar and beyond* (Varieties of English Around the World G39), 337–358. Amsterdam: John Benjamins. DOI: 10.1075/veaw.g39.19mul.

Nelson, Gerald. 2002a. *International Corpus of English: Markup manual for spoken texts.* https://www.ice-corpora.uzh.ch/dam/jcr:72c70d5a-8da8-496f-b8dc-5fb66986c87c/spoken (7 December, 2021).

Nelson, Gerald. 2002b. *International Corpus of English: Markup manual for written texts.* https://www.ice-corpora.uzh.ch/dam/jcr:df7b1e8f-005f-4346-903b-c77b6c1da66a/written (7 December, 2021).

Nelson, Gerald, Sean Wallis & Bas Aarts. 2002. *Exploring natural language: Working with the British component of the International Corpus of English* (Varieties of English Around the World G29). Amsterdam: John Benjamins. DOI: 10.1075/veaw.g29.

Östman, Jan-Ola & Mirjam Fried. 2005. The cognitive grounding of Construction Grammar. In Jan-Ola Östman & Mirjam Fried (eds.), *Construction Grammars: Cognitive grounding and theoretical extensions* (Constructional Approaches to Language 3), 1–13. Amsterdam: John Benjamins. DOI: 10.1075/cal.3.01ost.

Oxford English Dictionary Online. 2021. Oxford. https://www.oed.com (7 December, 2021).

Pastor-Gómez, Iria. 2011. *The status and development of N+N sequences in Contemporary English noun phrases* (Linguistic Insights 126). Frankfurt am Main: Peter Lang. DOI: 10.3726/978-3-0351-0193-5.

Pennycook, Alastair. 1994. *The cultural politics of English as an international language* (Language In Social Life). London: Longman. DOI: 10.4324/9781315843605.

Phillips, Betty S. 2006. *Word frequency and lexical diffusion* (Palgrave Studies in Language History and Language Change). Basingstoke: Palgrave Macmillan. DOI: 10.1057/9780230286610.

Platt, John, Heidi Weber & Ho Mian Lian. 1984. *The new Englishes.* London: Routledge & Kegan Paul.

Quirk, Randolph. 1954. *The concessive relation in Old English poetry.* New Haven: Yale University Press.

Quirk, Randolph. 1985. The English language in a global context. In Randolph Quirk & Henry G. Widdowson (eds.), *English in the world. Teaching and learning the language and literatures*, 1–6. Cambridge: Cambridge University Press.

Quirk, Randolph, Sidney Greenbaum, Geoffrey Leech & Jan Svartvik. 1985. *A Comprehensive Grammar of the English Language.* London: Arnold.

R Core Team. 2021. *R: A language and environment for statistical computing.* Version 4.1.1. https://www.r-project.org (7 December, 2021).

Radford, Andrew. 1981. *Transformational syntax: A student's guide to Chomsky's Extended Standard Theory* (Cambridge Textbooks in Linguistics). Cambridge: Cambridge University Press.

Raudenbush, Stephen W. & Anthony S. Bryk. 2002. *Hierarchical linear models: Applications and data analysis methods* (Advanced Quantitative Techniques in the Social Sciences 1). 2nd edn. Thousand Oaks: Sage.

Rissanen, Matti. 2002. *Despite* or *notwithstanding*? On the development of concessive prepositions in English. In Andreas Fischer, Gunnel Tottie & Hans Martin Lehmann (eds.), *Text types and corpora: Studies in honour of Udo Fries*, 191–203. Tübingen: Narr.

RStudio Team. 2009–2021. *RStudio: Integrated development for R.* Version 1.4.1717. https://www.rstudio.com (7 December, 2021).

References

Rudolph, Elisabeth. 1996. *Contrast: Adversative and concessive relations and their expressions in English, German, Spanish, Portuguese on sentence and text level* (Research in Text Theory 23). Berlin: Mouton de Gruyter. DOI: 10.1515/9783110815856.

Sarkar, Deepayan. 2008. *Lattice: Multivariate data visualization with R* (Use R!). Berlin: Springer. DOI: 10.1007/978-0-387-75969-2.

Sarkar, Deepayan. 2021. *lattice: Trellis graphics for R.* R package version 0.20-44. https://cran.r-project.org/web/packages/lattice/lattice.pdf (7 December, 2021).

Sarkar, Deepayan & Felix Andrews. 2019. *latticeExtra: Extra graphical utilities based on lattice.* R package version 0.6-30. https://cran.r-project.org/web/packages/latticeExtra/latticeExtra.pdf (7 December, 2021).

Schlüter, Julia. 2009. The conditional subjunctive. In Günter Rohdenburg & Julia Schlüter (eds.), *One language, two grammars? Differences between British and American English* (Studies in English Language), 277–305. Cambridge: Cambridge University Press. DOI: 10.1017/CBO9780511551970.016.

Schneider, Edgar W. 2003. The dynamics of new Englishes: From identity construction to dialect birth. *Language* 79(2). 233–281. DOI: 10.1353/lan.2003.0136.

Schneider, Edgar W. 2007. *Postcolonial English: Varieties around the world* (Cambridge Approaches to Language Contact). Cambridge: Cambridge University Press. DOI: 10.1017/CBO9780511618901.

Schneider, Edgar W. 2011. *English around the world: An introduction* (Cambridge Introductions to the English Language). Cambridge: Cambridge University Press. DOI: 10.1017/CBO9780511781711.

Schneider, Edgar W. 2014. New reflections on the evolutionary dynamics of World Englishes. *World Englishes* 33(1). 9–32. DOI: 10.1111/weng.12069.

Schützler, Ole. 2015. *A sociophonetic approach to Scottish Standard English* (Varieties of English Around the World G53). Amsterdam: John Benjamins. DOI: 10.1075/veaw.g53.

Schützler, Ole. 2017. A corpus-based study of concessive conjunctions in three L1-varieties of English. In Isabelle Buchstaller & Beat Siebenhaar (eds.), *Language variation – European perspectives VI: Selected papers from the Eighth International Conference on Language Variation in Europe (ICLaVE 8), Leipzig, May 2015* (Studies in Language Variation 19), 173–184. Amsterdam: John Benjamins. DOI: 10.1075/silv.19.11sch.

Schützler, Ole. 2018a. Concessive conjunctions in written American English: Diachronic and genre-related changes in frequency and semantics. In Richard J. Whitt (ed.), *Diachronic corpora, genre, and language change* (Studies in Corpus Linguistics 85), 195–218. Amsterdam: John Benjamins. DOI: 10.1075/scl.85.09sch.

Schützler, Ole. 2018b. *Concessive constructions in varieties of English*. University of Bamberg. (Postdoctoral thesis).

Schützler, Ole. 2018c. Grammaticalisation and information structure: Two perspectives on diachronic changes of *notwithstanding* in written American English. *English Language and Linguistics* 22(1). 101–122. DOI: 10 . 1017 / s1360674316000411.

Schützler, Ole. 2019. Variation and change at the interface of syntax and semantics: Concessive clauses in American English. In Claudia Claridge & Birte Bös (eds.), *Developments in English historical morpho-syntax* (Current Issues in Linguistic Theory 346), 247–268. Amsterdam: John Benjamins. DOI: 10.1075/cilt.346.12sch.

Schützler, Ole. 2020a. *Although*-constructions in varieties of English. *World Englishes* 39(3). 443–461. DOI: 10.1111/weng.12484.

Schützler, Ole. 2020b. Frequency changes and stylistic levelling of *though* in diachronic and synchronic varieties of English: Linguistic democratisation? *Language Sciences* 79(101266). DOI: 10.1016/j.langsci.2019.101266.

Schützler, Ole. 2021. *Concessive constructions in varieties of English: Corpus data.* Tromsø: DataverseNO. DOI: 10.18710/1JMFVR.

Schützler, Ole. 2022. Comparing methods for the evaluation of cluster structures in multidimensional analyses: Concessive constructions in varieties of English. In Ole Schützler & Julia Schlüter (eds.), *Data and methods in corpus linguistics: Comparative approaches*, 259–288. Cambridge: Cambridge University Press. DOI: 10.1017/9781108589314.010.

Schützler, Ole. 2023. Frequencies in corpus linguistics: Issues of scaling and visualization. In Lukas Sönning & Ole Schützler (eds.), *Data visualization in corpus linguistics: Critical reflections and future directions* (Studies in Variation, Contacts and Change in English 22). Helsinki: VARIENG. https://urn.fi/URN:NBN:fi:varieng:series-22-4.

Schützler, Ole, Ulrike Gut & Robert Fuchs. 2017. New perspectives on Scottish Standard English: Introducing the Scottish component of the International Corpus of English. In Joan C. Beal & Sylvie Hancil (eds.), *Perspectives on Northern English* (Topics in English Linguistics 96), 273–302. Berlin: Mouton de Gruyter. DOI: 10.1515/9783110450903-012.

Schützler, Ole & Jenny Herzky. 2021. Modal verbs of strong obligation in Scottish Standard English. *English Language and Linguistics* 26(1). 133–159. DOI: 10.1017/s1360674321000071.

Seidlhofer, Barbara. 2011. *Understanding English as a lingua franca*. Oxford: Oxford University Press.

References

Seoane, Elena & Cristina Suárez-Gómez (eds.). 2016. *World Englishes: New theoretical and methodological considerations* (Varieties of English Around the World G57). Amsterdam: John Benjamins. DOI: 10.1075/veaw.g57.

Shikano, Susumu. 2015. Bayesian estimation of regression models. In Henning Best & Christof Wolf (eds.), *The Sage Handbook of Regression Analysis and Causal Inference*, 31–54. London: Sage. DOI: 10.4135/9781446288146.

Smitterberg, Erik & Merja Kytö. 2015. English genres in diachronic corpus linguistics. In Philip Shaw, Britt Erman, Gunnel Melchers & Peter Sundkvist (eds.), *From clerks to corpora: Essays on the English language yesterday and today* (Stockholm English Studies 2), 117–133. Stockholm: Stockholm University Press. DOI: 10.16993/bab.g.

Snijders, Tom A. B. & Roel J. Bosker. 1999. *Multilevel analysis: An introduction to basic and advanced multilevel modelling*. London: Sage.

Sönning, Lukas & Julia Schlüter. 2022. Comparing standard reference corpora and Google Books Ngrams: Strengths, limitations and synergies in the contrastive study of variable *h-* in British and American English. In Ole Schützler & Julia Schlüter (eds.), *Data and methods in corpus linguistics: Comparative approaches*, 17–45. Cambridge: Cambridge University Press. DOI: 10.1017/9781108589314.002.

Sönning, Lukas. 2020. *Phonological variation in German learner English*. University of Bamberg. (Doctoral dissertation). DOI: 10.20378/irb-49135.

Speelman, Dirk, Kris Heylen & Dirk Geeraerts. 2018. Introduction. In Dirk Speelman, Kris Heylen & Dirk Geeraerts (eds.), *Mixed-effects regression models in linguistics* (Quantitative Methods in the Humanities and Social Sciences), 1–10. New York: Springer. DOI: 10.1007/978-3-319-69830-4_1.

Stan Development Team. 2011–2019. *Stan user's guide*. Version 2.27. https://mc-stan.org/docs/2_27/stan-users-guide/index.html (7 December, 2021).

Svartvik, Jan & Randolph Quirk (eds.). 1980. *A corpus of English conversation* (Lund Studies in English 56). Lund: Gleerup.

Sweetser, Eve E. 1990. *From etymology to pragmatics: Metaphorical and cultural aspects of semantic structure* (Cambridge Studies in Linguistics 54). Cambridge: Cambridge University Press. DOI: 10.1017/CBO9780511620904.

Thompson, Sandra A., Robert E. Longacre & Shin Ja J. Hwang. 2007. Adverbial clauses. In Timothy Shopen (ed.), *Language typology and syntactic description*, vol. 2: *Complex constructions*, 237–300. Cambridge: Cambridge University Press. DOI: 10.1017/CBO9780511619434.005.

Tizón-Couto, David & David Lorenz. 2021. Variables are valuable: Making a case for deductive modelling. *Linguistics* 59(5). 1279–1309. DOI: 10.1515/ling-2019-0050.

Tomasello, Michael. 2005. *Constructing a language: A usage-based theory of language acquisition*. Cambridge: Harvard University Press.

Tomasello, Michael. 2006. Construction Grammar for kids. *Constructions* Special Volume 1. 1–23. DOI: 10.24338/cons-452.

Traugott, Elizabeth C. 1989. On the rise of epistemic meanings in English: An example of subjectification in semantic change. *Language* 65(1). 31–55. DOI: 10.2307/414841.

Traugott, Elizabeth C. 2010. (Inter)subjectivity and (inter)subjectification: A reassessment. In Kristin Davidse, Lieven Vandelanotte & Hubert Cuyckens (eds.), *Subjectification, intersubjectification and grammaticalization* (Topics in English Linguistics 66), 29–71. Berlin: Mouton de Gruyter. DOI: 10.1515/9783110226102.1.29.

Traugott, Elizabeth C. 2014. Intersubjectification and clause periphery. In Lieselotte Brems, Lobke Ghesquière & Freek Van de Velde (eds.), *Intersubjectivity and intersubjectification in grammar and discourse: Theoretical and descriptive advances* (Benjamins Current Topics 65), 7–27. Amsterdam: John Benjamins. DOI: 10.1075/bct.65.02trau.

Trousdale, Graeme. 2012. Grammaticalization, constructions and the grammaticalization of constructions. In Kristin Davidse, Tine Breban, Liselotte Brems & Tanja Mortelmans (eds.), *Grammaticalization and language change: New reflections* (Studies in Language Companion Series 130), 167–198. Amsterdam: John Benjamins. DOI: 10.1075/slcs.130.07tro.

Tsunoda, Mie. 2012. Five-level classification of clause linkage in Japanese. *Studies in Language* 36(2). 382–429. DOI: 10.1075/sl.36.2.06tsu.

van der Auwera, Johan (ed.). 1998. *Adverbial constructions in the languages of Europe* (Empirical Approaches to Language Typology 20-3). Berlin: Mouton de Gruyter. DOI: 10.1515/9783110802610.

Verhagen, Arie. 2000. Concession implies causality, though in some other space. In Elizabeth Couper-Kuhlen & Bernd Kortmann (eds.), *Cause – condition – concession – contrast: Cognitive and discourse perspectives* (Topics in English Linguistics 33), 361–380. Berlin: Mouton de Gruyter. DOI: 10.1515/9783110219043-015.

Vetter, Fabian. 2021. *Issues of corpus comparability and register variation in the International Corpus of English: Theories and computer applications*. University of Bamberg. (Doctoral dissertation). DOI: 10.20378/irb-52406.

Vetter, Fabian. 2022. Comparing approaches to (sub-)register variation: The "press editorials" sections in the British, Canadian and Jamaican components of ICE. In Ole Schützler & Julia Schlüter (eds.), *Data and methods in corpus lin-*

guistics: Comparative approaches, 75–100. Cambridge: Cambridge University Press. DOI: 10.1017/9781108589314.004.

Werner, Valentin. 2014. *The present perfect in World Englishes: Charting unity and diversity* (Bamberger Beiträge zur Linguistik 5). Bamberg: University of Bamberg Press.

Wiechmann, Daniel & Elma Kerz. 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). 1–23. DOI: 10.1017/S1360674312000305.

Winter, Bodo. 2020. *Statistics for linguistics: An introduction using R.* New York: Routledge. DOI: 10.4324/9781315165547.

Winter, Bodo & Martine Grice. 2021. Independence and generalizability in linguistics. *Linguistics* 59(5). 1251–1277. DOI: 10.1515/ling-2019-0049.

Wunder, Eva-Maria, Holger Voormann & Ulrike Gut. 2010. The ICE Nigeria corpus project: Creating an open, rich and accurate corpus. *ICAME Journal* 34. 78–88.

Yáñez-Bouza, Nuria. 2011. ARCHER past and present (1990–2010). *ICAME Journal* 35. 205–236.

Name index

Aarts, Bas, 3, 5, 9, 11, 13, 31, 33, 38, 51, 77, 78, 83, 84, 86, 88, 91, 94, 126, 175, 176, 203, 204, 207
Abercrombie, David, 70
Afifi, Abdelmonem, 222
Agresti, Alan, 100, 103, 105–109, 111, 112
Akinnaso, F. Niyi, 65, 66
Altenberg, Bengt, 3, 28, 76–79, 83, 84, 86, 88, 126, 148, 203, 204, 210
Andrews, Felix, 112
Anscombre, Jean-Claude, 18
Anthony, Laurence, 96
Azar, Moshe, 2, 18, 19, 80

Baayen, R. Harald, 99, 101, 112
Baker, Paul, 35
Barth-Weingarten, Dagmar, 1
Bates, Douglas, 103
Benveniste, Émile, 17
Bergs, Alexander, 56, 57
Berlage, Eva, 3, 5
Bernaisch, Tobias, 94
Best, Henning, 107, 111
Bhatt, Rakesh M., 67, 69, 71
Biber, Douglas, 4, 9, 31, 65, 66, 75, 76, 83, 84, 86, 88, 126, 175, 204, 207
Bolstad, William M., 102, 112
Borkin, Ann, 14
Bosker, Roel J., 100, 107

Brinton, Donna M., 29
Brinton, Laurel J., 29
Bryant, Margaret M., 14
Bryk, Anthony S., 105
Bürkner, Paul-Christian, 103
Burnham, Josephine M., 4, 11, 14, 18, 21, 75, 215
Buschfeld, Sarah, 71, 72
Bybee, Joan L., 55, 58–60

Calin-Jageman, Robert, 112
Cameron, A. Colin, 106
Cappelle, Bert, 57
Chafe, Wallace L., 27–29, 64–66, 95, 96
Chen, Guohua, 14
Clyne, Michael, 67
Conrad, Susan, 66
Couper-Kuhlen, Elizabeth, 1, 2
Crawford, William J., 30
Crevels, Mily, 15, 17, 18, 20, 21, 27
Croft, William, 55, 56
Cruse, D. Alan, 55
Culicover, Peter W., 31
Cumming, Geoff, 112
Curran, James M., 102, 112

Danielewicz, Jane, 65
Davies, Mark, 9, 35, 78, 80, 85
Di Meola, Claudio, 2, 4, 11, 13, 19, 21, 23
Diessel, Holger, 27–30, 82, 83, 90

Name index

Language index

Subject index

Concessive constructions in varieties of English

This volume presents a synchronic investigation of concessive constructions in nine varieties of English, based on data from the *International Corpus of English*. The structures of interest are complex sentences with a subordinate clause introduced by *although*, *though* or *even though*. Various functional and formal features are taken into account: (i) the semantic/pragmatic relation that holds between the propositions involved, (ii) the position of the subordinate clause, (iii) the conjunction that is used, and (iv) the syntax of the subordinate clause. By exploring patterns of variation from a Construction Grammar perspective, the study works towards an explanatory model whose point of departure is at the functional (semantic/pragmatic) level and which makes hierarchically organised predictions for different formal levels (clause position, choice of connective and realisation of the subordinate clause). It treats concessives as complex form-function pairings and develops arguments and routines that may inform quantitative approaches to constructional variation more generally.