BRILL

COGNITIVE SEMANTICS 10 (2024) 158–189

COGNITIVE SEMANTICS
brill.com/cose

If Not For-If It Weren't/Wasn't For Counterfactual Constructions: A Multivariate Extension of Collostructional Analysis

Jesús Olguín Martínez
Department of English, Illinois State University, Normal, Illinois, US
olguinmartinez@ucsb.edu

Stefan Th. Gries
Department of Linguistics, UC Santa Barbara, California, USA &
Department of English, JLU Giessen, Germany
stgries@gmail.com

Received 23 January 2024 | Accepted 6 June 2024 |
Published online 4 October 2024

Abstract

The present paper investigates *if not for-if it weren't/wasn't for* constructions in American English based on the analysis of more than 2,100 constructions extracted from The Corpus of Contemporary American English (COCA). In analyzing these constructions, we focus on significant interdependencies between the slots of the protasis (i.e. animacy of the referent of the NP) and apodosis (i.e. semantics of the verb lemma) and how these cross-clausal associations interact with other linguistic variables, such as the polarity of the apodosis. We apply a new multivariate extension of collostructional analysis that combines distinctive and (co-varying) collexeme analysis via a hierarchical configural frequency analysis. This allows the analyst to identify associations not just of one slot to a construction or one slot to one other slot in one construction, but to include other features to identify (i) which constructions are preferred by which fillers in, now, one or more slots of one construction and (ii) which constructions are preferred by which (combinations) of one or more fillers.

 | DOI:10.1163/23526416-BJA10067
 | ISSN: 2352-6408 (print) 2352-6416 (online)

Keywords

constructions – semantics – counterfactuals – corpus data – collostructional analysis – hierarchical configural frequency analysis

1 Introduction: Setting the Stage

People have the ability to specify alternatives that are in some way better or worse than actuality. This ability is implicated in diverse cognitive activities – including daydreams, fantasy, deduction, and calculating probabilities – and provides "the building blocks for generating imaginary possibilities in creative and insightful ways" (Byrne, 2002: 426). Such mental acts of imagining what might have been are often linguistically encoded in counterfactual constructions (e.g. *if you had gone, you would have seen her*) and compare reality with what might have been (Declerck & Reed, 2001: 13). Such constructions include the kinds of construction exemplified in (1), but may also show up in other guises as well, as in (2), (3), and (4).

(1) *If I had known that, I wouldn't have appointed him.*
(2) *The child is crying, as if I had hit him.*
(3) *I should have done it!*
(4) *If only she had come!*

One counterfactual construction that has been traditionally neglected in the corpus-based and typological literature is the *if not for* constructions as in (5):

(5) *If not for my protector Bane, I would have died.*

For example, Klomps (2021) analyzes the differences in verb tense usage of Dutch conditional constructions, including counterfactual conditionals, but does not take into account *if not for* constructions. In another corpus-based study, Declerck & Reed (2001) analyze conditional constructions in English focusing especially on Tense-Aspect-Mood (TAM) values of different types of counterfactual conditionals (including counteridenticals like *if I were you, I would have paid on time*), but they, too, do not include *if not for* constructions (Declerck & Reed, 2001: 177). This counterfactual construction has also been neglected in other corpus-based studies of other languages, such as Mandarin

(Jing-Schmid, 2017) as well as in typological studies (e.g. Olguín Martínez & Lester, 2021).

The present investigation seeks to contribute to fill this gap with a corpus-based analysis of *if not for* constructions in contemporary American English. They involve a reversal of polarity in that clauses with positive polarity include a corresponding negative proposition in their interpretation whereas clauses with negative polarity include a corresponding positive proposition (Van linden & Verstraete, 2008: 4). In (6), the apodosis involves a negative polarity and implies that Michael did it. On the other hand, in (7), the apodosis involves a positive polarity and implies that Carol did not die.[1]

(6) [$_{\text{protasis}}$ *If not for them*], [$_{\text{apodosis}}$ *Michael wouldn't have done it*] (*I am glad he did it*)].
(7) [$_{\text{protasis}}$ *If not for Rachel*], [$_{\text{apodosis}}$ *Carol would have died* (*it's good she didn't die*)].

If not for constructions bear resemblance to other constructions such as *if it weren't for* (8) and *if it wasn't for* constructions (9). These pairs of functionally – meaning semantically, discourse-functionally, or information-structurally – more-or-less equivalent expressions are frequent and widespread in English.

(8) *If it weren't for foolish decisions I made, none of this would have happened.*
(9) *I never even would have thought about doing this if it wasn't for you.*

Here we explore the category of *if not for-if it weren't/wasn't for* constructions in American English based on more than 2100 constructions extracted from The Corpus of Contemporary American English (COCA). The following *if not for* and *if it weren't/wasn't for* constructions are considered here:

(10) [$_{\text{protasis}}$ *If not for*___], [$_{\text{apodosis}}$ NP *would/may/might* (NEG) (*have*)___].
(11) [$_{\text{protasis}}$ *If it weren't for*___], [$_{\text{apodosis}}$ NP *would/may/might* (NEG) (*have*)___].
(12) [$_{\text{protasis}}$ *If it wasn't for*___], [$_{\text{apodosis}}$ NP *would/may/might* (NEG) (*have*)___].

As can be seen in the constructions shown above, protases occur with Noun Phrases (NPs) that may be human, animate, or inanimate, while apodoses may

1 We use *protasis* and *apodosis* to refer to the *if not for* clause and the main clause respectively.

appear with different types of verb lemmas, and may show positive or negative polarity.

Given that *if not for-if it weren't/wasn't for* constructions are unexplored territory in both corpus-based and typological studies, the present study explores these constructions in a cognitive/usage-based and collostructional way.[2] Let us next outline the expectations, a word we are using to avoid the word *hypothesis* that might trigger a hypothesis-testing approach – this study, like most collostructional studies, is exploratory and descriptive, not statistically inferential. Different types of counterfactual conditionals can have a benefactive or a malefactive sense. *If not for-if it weren't/wasn't for* are not an exception in that they can have a benefactive (13) or a malefactive meaning (14).

(13) *I would have been killed, if not for him* (*thanks to him, I was not killed*).
(14) *If not for you, I would be successful* (*because of you, I wasn't successful*).

We assume that benefactive *if not for-if it weren't/wasn't for* constructions mean 'due to/thanks to X's support/help/presence, Y achieved or avoided Z.' In the absence of X (benefactor), Y (benefactum) would have failed to accomplish Z or would not have avoided Z. Note that benefaction may be more direct or indirect in nature (Kittilä & Zúñiga, 2010: 2). On the other hand, malefactive *if not for-if it weren't/wasn't for* constructions mean 'in the absence of X, Y would not have failed to accomplish Z or would have been able to avoid an unpleasant situation.' X (maleficiary) represents a volitional causation that affects Y (affectee) adversely. Put another way, Y is affected negatively by X in a direct fashion.

Given the semantics of benefactives, we expect they will occur with protases with humans, animates, or inanimates when the apodoses appear with detriment verbs (e.g. *kill, die*) and show positive polarity, as in (15). Moreover, we expect that they will occur with the same kinds of protases when the apodoses

2 One way in which corpus-based approaches have studied associative connections between individual lexemes and specific slots of constructions is by means of what has become known as methods from the family of collostructional analyses, which comes in three 'flavors':
- collexeme analysis, which quantifies to what extent word types in one slot of a construction 'like to occur' in that construction (Stefanowitsch & Gries, 2003);
- distinctive collexeme analysis, which quantifies to what extent word types in one slot of two or more functionally similar constructions 'prefer to occur' in one or more of the constructions (Gries & Stefanowitsch, 2004a);
- co-varying collexeme analysis, which quantifies to what extent word types in two slots of one construction 'like to co-occur' (Gries & Stefanowitsch, 2004b).

appear with accomplishment verbs (e.g. *have, get, become*) and show negative polarity, as in (16).[3]

(15) *If not for her, I would have died.*
(16) *If not for her, my dreams wouldn't have become true.*

Malefactives, on the other hand, are expected to occur with protases with humans when the apodoses appear with accomplishment verbs (e.g. *have, get, become*), and show positive polarity, as in (17). We also expect that protases will occur with humans when the apodoses appear with state verbs (e.g. *be*) and show negative polarity, as in (18).

(17) *If not for Andrew, I would have a better job.*
(18) *If not for you, I wouldn't be in jail.*

Table 1 provides a summary of our expectations. The CxType column contains information on whether the construction is benefactive or malefactive and the other columns contain information regarding the expected meaning/semantics of each construction and information regarding the preference patterns we expect for benefactive and malefactive *if not for-if it weren't/wasn't for*. Note that in the apodosis lemma column, we list the verbs we expect will occur in each cluster followed by, in parentheses, verb categories. For instance, the verb lemmas *kill* and *die* are lemmas that are categorized as detriment verbs in the present study.

Theoretically, we adopt a usage-based Construction Grammar approach to explore the expectations outlined in Table 1. Thus, we assume that there are associative connections between individual lexemes and specific slots of constructions. Put another way, linguistic structure is lexically particular in the sense that constructions are generally associated with specific lexical items, which gives rise to a joint distribution of lexemes in constructions that are known in the literature as *filler-slot relations* (Diessel, 2019: 20). It has been shown that filler-slot relations must be analyzed in terms of matching categories (Fillmore & Kay, 1999), which for our usage-based perspective means

3 For readability's sake, we will henceforth write "humans", "animates", or "inanimates" to avoid the reading flow-disrupting "human referents of NPs", "animate referents of NPs", or "inanimate referents of NPs"; we will likewise just write "*be* prefers …" as opposed to "the verb lemma *be* prefers …" and "benefactives" and "malefactives" as opposed to "benefactive constructions" and "malefactive constructions" respectively.

TABLE 1 Summary of expectations

(Dis) Pref	Construction	Cx Type	NP Type	Apodosis Lemma	Polarity
pref	any	benefactive	any	*kill, die* (detriment)	positive
pref	any	benefactive	any	*have, get, become* (accomplishment)	negative
pref	any	malefactive	human	*have, get, become* (accomplishment)	positive
pref	any	malefactive	human	*be* (state)	negative

that the distributional biases and constraints of lexemes and constructions are not arbitrary, but functionally motivated and predictable. This is in essence Goldberg's (1995: 50) *semantic coherence principle*, according to which constructions attract lexical items that are compatible with the semantic specifications of particular slots. Usage-based construction grammarians argue there are always multiple ways of viewing the same experience so that speakers must choose from a range of linguistic means to construe the event in a certain way (Croft & Cruse, 2004: §3). This means that the *conceptualization* of the same or similar experience may become associated with particular lexemes and constructions (Diessel, 2019: 29).

Methodologically, unlike the vast majority of collostructional analyses, we do not restrict our attention to just one slot in each construction (e.g. the verb slot), but we consider multiple linguistic features at the same time, which amounts to making collostructional analysis multivariate. Specifically, we employ a new multivariate extension of collostructional analysis that combines distinctive and (co-varying) collexeme analysis via a hierarchical configural frequency analysis. This allows the analyst to identify associations not just of one slot to a construction or one slot to one other slot in one construction, but to include as many other features as one can reasonably accommodate to identify (i) which constructions are preferred by which fillers in, now, one or more slots of one construction and (ii) which constructions are preferred by other characteristics of the construction or in its context. We use the three counterfactual constructions as a proof-of-concept kind of application for our multivariate extension.

Admittedly, other attempts have been made to enhance collexeme analysis in a multivariate way. However, their focus has been on fewer dimensions than

the ones we consider here. Accordingly, our method provides a considerably higher degree of descriptive accuracy than previous attempts. For instance, Stefanowitsch & Flach (2020) explore the combination of a distinctive and a co-varying collexeme analysis for studying the *too* ADJ *to* V and ADJ *enough to* V constructions, meaning they involve only three dimensions of information – columns in their data frames – in their analysis.

Hoffmann et al. (2019) involve a similar number of dimensions in their demonstration of the value of exploring preferential patterns of how slots in constructions consisting of two clauses are filled (such as *the X, the Xer* comparative correlative as in *the older I get, the happier I am*). In other words, they extend collostructional methods by exploring the elements in slots on a more schematic level and the correlations between what happens in a construction's slots on that more abstract level and show that in English comparative correlative constructions, the grammatical/phrasal filler types of both clauses (e.g. adjectives, adverbs, noun phrases) interact with other construction's characteristics, such as lexical fillers (e.g. 'the more … the more …'; 'the less, ….the less..'), and the presence/absence of different kinds of deletion.

Finally, a higher number of dimensions is included by Hampe & Gries (2018), who bring a multifactorial predictive-modeling perspective to the study of meso-/mini-constructions, i.e. constructions that are sub-constructions of a more general/schematic construction, or allostructions of a constructeme. They annotate 560 corpus instances of eight meso-constructions resulting from the combination of four temporal subordinators (*after*, *before*, *once*, and *until*) and two clause orders (matrix clause before adverbial clause and the other way round) and annotate them for four binary predictors (type of matrix clause, verb type in matrix clause, presence of negative polarity markers, and presence of perfects in adverbial clause). They then use a multinomial regression to (i) determine whether the stipulation of the eight meso-constructions is supported by a high degree of predictive power of the regression model and (ii) identify multifactorially-defined prototypes of the eight meso-constructions.

The remainder of the paper is structured as follows. In Section 2, the corpus data are presented and the variables for which all *if not for-if it weren't/wasn't for* constructions have been annotated are discussed. We also elaborate on the statistical approach used here to analyze these complex sentence constructions. In Section 3, we present the results of the multivariate extension of collostructional analysis and, on the basis of these results, Section 4 will propose that the different degrees of resolution on the preference patterns characterizing *if not for-if it weren't/wasn't for* constructions can be explained by different semantic factors. Section 5 offers some concluding remarks.

2 Methods

This section introduces the corpus data and outlines our multivariate extension of collostructional analysis. Collostructional analysis is a family of methods grounded in two frameworks, one theoretical and one methodological. The theoretical framework is provided by usage-based constructional theories of grammar. The methodological framework is that of quantitative corpus linguistics and is characterized by three features: (a) it is based on natural discourse data from representative and balanced corpora, (b) the linguistic phenomenon in question is retrieved exhaustively from the corpus, (c) the data are subjected to strict quantification and statistical evaluation (Gries & Stefanowitsch, 2010: 74).

2.1 *Corpus Data, Data Extraction, and Annotation*

The English corpus chosen for the investigation of *if not for-if it weren't/wasn't for* constructions in American English was COCA (Davis, 2008). The data used here consist for the most part of web-based material (e.g. newspapers) and material from TV and movies subtitles.

The procedure for data retrieval was the following. A near-exhaustive concordance of *if not for-if it weren't/wasn't for* constructions was performed by searching the COCA for the forms *if not for, if it weren't for*, and *if it wasn't for*. This generated a large sample of constructions occurring with the lexical items *if not for, if it weren't for*, and *if it wasn't for*, which was then trimmed down to exclude other constructions that are not relevant to the present study, such as the following:

(19) *If not for* as 'quite possibly' (e.g. *he suffered for days if not for years*).
(20) *If not for* as 'either … or …' (e.g. *do it if not for his sake for the sake of your child*).
(21) *If not for* as 'only/except that' (e.g. *in fact he was quite comely if not for the ridiculous wig on the top of his head*).
(22) *If not for* as 'or else' (e.g. *put the milk on the fridge, if not, it will go bad*).

The resulting dataset contains 2156 instances of the variable CONSTRUCTION with the three levels *if not for, if it weren't for*, and *if it wasn't for.* For these constructions, we then also coded the relevant variables for our analysis:

– CXTYPE: whether *if not for-if it weren't/wasn't for* constructions are benefactive or malefactive;
– NPTYPE: type of noun phrase that occurs in the protasis;

– ApodosisLemma: verbs that can occur in the slot of the apodosis;
– Polarity: whether the apodosis shows positive or negative polarity.

These were manually annotated by inspecting each of the 2156 *if not for*, *if it weren't for*, and *if it wasn't for* constructions. Table 2 shows the way we organized our data and the subsequent sections describe the variable levels and their annotation in more detail.

2.1.1 Type of Construction: Benefaction and Malefaction

As mentioned in Section 1, *if not for-if it weren't/wasn't for* constructions can have a benefactive or a malefactive meaning. This fact has not gone unnoticed in experimental and corpus-based studies and is often referred to as *semantic prosody*. It has been argued that semantic prosodies have mental representation and are stored as part of a template, i.e. a construction (Smith & Nordquist, 2012: 293), and that would mean that a semantic prosody is not part of a word's lexical semantics, as has been argued by different linguists (e.g. Louw, 1993) – rather, it has a discourse property (Morley & Partington, 2009).

An example often cited in the literature is that of *cause*. In different experiments (e.g. Nordquist, 2004), it has been shown that *causation* per se is neither benefactive nor malefactive. In other words, there is nothing about the form or meaning of the lexeme *cause* alone that would allow us to predict its benefactive or malefactive interpretation. Instead, *cause* can receive a benefactive (as in (23)) or a malefactive meaning (as in (24)) depending on the discourse context.

(23) *a cause for joy/celebration/merriment.*
(24) *a cause for tears.*

The situation is similar for *if not for-if it weren't/wasn't for* constructions, whose semantic prosody also resides at the discourse level (as opposed to being part of either construction's semantics), as is exemplified in the benefactive case of (25) or the malefactive case of (26). However, since frequency affects the activation and processing of lexemes and constructions, which can have long-lasting effects on the development of linguistic structure (Divjak, 2019), over time certain semantic prosodies could end up being entrenched as part of the core semantics of *if not for-if it weren't/wasn't for* constructions, namely when particular lexical configurations are frequently used with the same or similar semantic prosodies.

TABLE 2 Organization of the *if not for-if it weren't/wasn't for* data in the present study

Source	Example	Cx Type	NP Type	Apodosis Lemma	Polarity
awards-daily.com	*if not for her, I wouldn't be successful*	benefactive	human	*be* (state)	negative
cbsnews.com	*if not for them, we would have more money*	malefactive	human	*have* (accomplishment)	positive
Look Who's Talking Now	*if not for his father, he would know how to repair cars*	malefactive	human	*become* (accomplishment)	positive
Fox: First 100 Days	*And it might be worse, if not for Ted Frank*	benefactive	human	*be* (state)	positive

(25) *I am thankful that my father gave me that job. I don't complain about it! If not for him, I wouldn't have been able to pay my bills on time. Thanks to him I didn't go to jail.*

(26) *I am sad I didn't play the final. If not for her, I would've been able to play the final match. I really hate her.*

Determining whether an *if not for*, *if it weren't for*, or *if it wasn't for* construction was benefactive or malefactive was one of the most arduous parts of the data analysis. For this task, it was not sufficient to analyze biclausal constructions as in (27) without taking into account their discourse context. For instance, without taking into account the discourse context of (27), it is not clear whether this

construction should be characterized as involving a benefactive or malefactive meaning. However, by looking at preceding and/or subsequent stretches of discourse of the biclausal constructions, it was possible to determine that the example in (27) is attested in a malefactive discourse context, as can be seen in (28). It is this methodological step that enabled us to determine whether a construction was benefactive or malefactive.

(27) *If not for you, I wouldn't be here.*

(28) *It was Michael who stole the bank and not me. I am disappointed he didn't confess it.* ***If not for him, I wouldn't be here****. Jail is horrible. You sit around all day with nothing to do but work out.*

2.1.2 Type of NP in the Protasis

Animacy has been shown to interact with many linguistic and extra-linguistic features in intriguing ways, such as case and agreement, word order, form of reference (including pronoun selection), inverse alignment and affixation, argument structure patterns, thematic structure, and noun class systems (Silverstein, 1976). Accordingly, the role of animacy in the grammars of the world's languages is extremely diverse.

Given that animacy is often seen as a functional primitive which plays a part in structuring human grammars (Dahl, 2008), it seems reasonable to assume that the animacy of the protasis NP of *if not for-if it weren't/wasn't for* constructions will interact with other features of these constructions (see Table 1).

For the purpose of the present study, we decided to adopt a tripartite distinction of animacy and classify protasis NPs as human, animate, and inanimate. Note that there were a number of examples in which it was difficult to determine the animacy of a protasis NP. In particular, demonstratives represented a challenging task. In the example in (29), the protasis NP is the demonstrative *that* and, at first glance, it is not clear how it should be classified. In this scenario, too, it was necessary to analyze the preceding stretches of discourse of these biclausal constructions to determine the animacy of demonstratives.

(29) *If not for that, I would have gone.*

2.1.3 Apodosis Lemma

We annotated each construction with regard to the lemma occurring in the apodosis of *if not for-if it weren't/wasn't for* constructions, but there also were examples of apodoses with two or more verbs, as in (30) and (31):

(30) *If not for him, I would know how to do it.*
(31) *If not for them, I wouldn't want to buy a car.*

In such examples, apodoses are complement clauses. The predicate *know* in (30) entails reference to another proposition (i.e. *how to do it*). In a similar fashion in (31), the predicate *want* denotes a mental activity that is inherently directed at, and hence entails reference to, another proposition (i.e. *to buy a car*). These examples involve a complement-taking predicate (i.e. *know* and *want*) and a complement clause verb (i.e. *do* and *buy*). For these examples, we only took into account the complement-taking predicate.

2.1.4 Apodosis Polarity

As to the polarity of the apodosis of *if not for-if it weren't/wasn't for* constructions, sometimes they appear with different polarity as in (32) or (33).

(32) *If not for him, I wouldn't be here, but I am glad I am here.*
(33) *If not for my sister, I would have been happier last night at the meeting, but I wasn't!*

As discussed in Section 1, *if not for-if it weren't/wasn't for* constructions involve a reversal of polarity in that clauses with positive polarity include a corresponding negative proposition in their interpretation, and clauses with negative polarity include a corresponding positive proposition. In most constructions in the dataset, such reversal of polarity is inferred. However, there are cases in which it is explicitly indicated by a 'but' clause, as in (32) and (33). For these examples, we decided to take into account the polarity of the apodosis and not that of the 'but' clause. On these grounds, the polarity of the apodosis in the example in (32) is negative and the polarity of the apodosis in (33) is positive.

2.2 *The Statistical Analysis*

As mentioned in Section 1, a methodological goal of this paper is to increase the resolution of collostructional methods. While collostructional methods have always been based on the Distributional Hypothesis – the notion that distributional similarity/patterns reflect functional similarity/patterns – and have yielded an enormous amount of insightful results regarding constructions' functions, it is probably also accurate to state that these methods have underutilized the rich distributional information corpora have to offer. This is because collostructional methods as they have been used for 20 years are mostly univariate: collexeme analyses looked at one variable slot (words, often

verbs) in one construction, distinctive collexeme analyses looked at one comparable variable slot (words, often verbs) in two or more constructions, and co-varying collexeme analyses looked at the combination of two slots in one construction, and then the distributional behavior in that one slot, or the combination of the two slots in co-varying collexeme analysis, was then quantified with some association measure (often, but not necessarily, the (log-transformed) *p*-value of a Fisher-Yates exact test).

In this study, we increase the focus by including all the annotated features described above at the same time and applying a hierarchical configural frequency analysis (HCFA) to coin one of two recently proposed extensions of collostructional methods which we call multivariate collexeme analysis. HCFA is a multivariate extension of co-occurrence frequency tables (see Gries, 2009: Section 5.1) but has so far only been applied to collostructional studies very rarely (and with only two or maximally three features, see Stefanowitsch & Gries, 2005; Newman, 2011). Given a set of *f* annotated features (including the chosen construction itself), the HCFA algorithm generates all possible frequency tables of all possible subsets of the *f* features. For example, if *f*=3 with the features called A, B, and C, a HCFA generates the following seven tables:

- the frequencies of the levels of A, B, and C;
- the frequencies of the levels of A:B (i.e. A cross-tabulated with B), A:C, and B:C;
- the frequencies of the levels of A:B:C (i.e. A cross-tabulated with B and C).

For each of the tables, the HCFA determines:

- the observed co-occurrence frequencies;
- the expected co-occurrence frequencies (under the null hypothesis of complete independence of features);
- contributions to chi-squared and to *G*-squared;
- an unadjusted two-tailed *p*-value of an exact binomial test comparing the observed frequency against the expected one, and a two-tailed *p*-value of an exact binomial test that has been adjusted for as many post-hoc tests (using Holm's sequential method to control the family-wise error rate) as the relevant co-occurrence table had cells; for this more exploratory application, we use a *p*-value threshold of 0.1, *but* for the *p*-values adjusted for multiple post-hoc tests (something that few collostructional studies have done);
- a measure called the Q coefficient that quantifies the discrepancy of observed and expected frequency with a score falling into the interval [0, 1].

Two aspects of an HCFA are particularly noteworthy in the present context. First, because it generates all possible sub-tables as described above, the analysis can identify co-occurrence preferences (or dispreferences) that go way beyond a single slot and a construction – rather, it can, for instance, see whether a certain construction is preferred with certain verbs in its main slot *and* certain animacy values in one other slot, *but only* when the polarity is positive and the construction is malefactive, which provides a much higher degree of descriptive accuracy. Second, because of the formation of all sub-tables, one can also determine which (dis)preferences observed in some higher-dimensional sub-table are actually not required because they do not meaningfully or significantly qualify the distinctions of a lower-dimensional sub-table. For instance, if the sub-table for features A:B:C does not provide information that is different from what sub-tables A:B and B:C already provide, the analyst can infer from that the constructional patterns in A:B and B:C are not further specified, or remain schematic, and that the more precise resolution of A:B:C is not required. This way, using argumentation reminiscent of Occam's razor, A:B:C can be said to not contain distributional information that is required for understanding the construction's behavior.

In this study we applied the HCFA to the '5-dimensional space' covered by our variables CONSTRUCTION, CXTYPE, NPTYPE, APODOSISLEMMA, and POLARITY. To that end, we used an R script HCFA written by the second author, which is an update of the script used in the above-cited previous studies.

3 Results

Because of the multivariate and hierarchical nature of the approach, the amount of results it generates is very voluminous. To tackle the many results while still remaining readable, the results are organized by construction and, within each construction, by which configurations of features are preferred (i.e. occur more often than expected by chance, which in HCFA literature are referred to as *types*) or dispreferred (i.e. occur less often than expected by chance, which in HCFA literature are referred to as *antitypes*) with $p_{adjusted}<0.1$,[4] and we organize the discussion into related clusters of types and antitypes where and as much as possible (mostly on the basis of the values of APODOSISLEMMA together with one other variable).

4 The full results spreadsheet can be accessed at the second author's website at https://www.stgries.info/research/2024_JFOM-STG_CounterfacturalsMultVarColl_CognSem.ods.

3.1 *If Not for Constructions*

3.1.1 Preferred Co-occurrences

One cluster *if not for* 'likes' is a group of verbs that one might classify as 'having to do with damage, disadvantage, or injury' and that occur in positive polarity contexts, such as:

– *lose* and *hurt* (as the ApodosisLemma) especially with benefactives (for CxType), as in (34) and (35);
– *fail* and *kill* especially so with humans (for NPType) and benefactives (for CxType), as in (36) and (37).

(34) *If not for them, I would have lost the final match.*
(35) *If not for her, I would have hurt myself with the knife.*
(36) *If not for him, I would have failed the test. I am glad he helped me to study for the test. My mom would have gotten upset if I hadn't passed math.*
(37) *If not for her, she would have killed Susan.*

The *if not for* construction also likes *seem* with positive polarity (especially with inanimates and benefactives), as in (38), and *die* with the combination of humans and benefactives, as in (39).

(38) *If not for her makeup, Laura would have seemed sicker.*
(39) *If not for him, they would have died.*

On the other hand, for negative polarity contexts, the *if not for* construction occurs with *be*, especially with humans and malefactives, as in (40). In negative polarity contexts, this counterfactual construction also likes *be able to*, especially with humans, as in (41). The last two clusters include cases in which *if not for* constructions with negative polarity appear with *exist* (especially with benefactives and humans or inanimates) as in (42) and *know* (especially with benefactives and inanimates) as in (43).

(40) *If not for her, I wouldn't be here in this bad place.*
(41) *If not for Mr. Taylor, I wouldn't be able to pay my bills on time.*
(42) *If not for my parents, I wouldn't exist.*
(43) *If not for this TV commercial, my father wouldn't know the negative effects of alcohol.*

3.1.2 Dispreferred Co-occurrences

The antitypes of *if not for* are fairly diverse and suggest fewer interpretively useful patterns. With negative polarity, the construction dislikes malefactives (especially with *have* and humans or inanimates) as in (44); also, *if not for* dislikes negative polarity with *die* (especially with benefactives or with humans), as in (45).

(44) *If not for Jacob, her mother wouldn't have this pain in her chest. He's always acting like a stupid child!*
(45) *If not for Dr. Morales, my mother wouldn't have died. We are very grateful because Dr. Morales found out a way to help her see God. After years of fighting cancer, she can now rest in peace.*

If not for constructions showing positive polarity dislike *be able to* or *know* with humans or benefactives, as in (46) and (47) respectively. However, as for *know*, it is worth noting that *if not for* constructions dislike this verb lemma with malefactives in general, as in (48).

(46) *If not for her aunt, Katherine would be able to spend money on useless stuff. I thank God she taught her how to spend her money wisely.*
(47) *If not for my sister, I would know how to rob banks. Thanks to her, I was able to stay away from bad friendships.*
(48) *If not for him, I would know what to do in this horrible situation.*

Other *if not for* antitypes showing positive polarity are the following:

– *have* with benefactives (49):

(49) *If not for him, I would have worse grades at school.*

– Inanimates with malefactives (50):

(50) *If not for this parking lot, I would be in a better place.*

3.2 *If it weren't for Constructions*

3.2.1 Preferred Co-occurrences

With one exception – benefactives with *have* and negative polarity (51) – this counterfactual conditional construction much prefers malefactives and positive polarity contexts, as in (52).

(51) *If it weren't for Charles, I wouldn't have what I have today.*
(52) *If it weren't for me, you would be in a better school district.*

Specifically, *if it weren't for* likes malefactives with positive polarity, especially with *know* (53), but also with humans and *get* (54); finally, it likes malefactives with humans or inanimates, as in (55).

(53) *If it weren't for Nate, Michael would have known how to behave there.*
(54) *If it weren't for you, I would have gotten a better job.*
(55) *If it weren't for the new medicine, I would feel healthier.*

3.2.2 Dispreferred Co-occurrences

This construction has very general dislikes/antitypes: it rejects negative polarity in general as well as benefactives in general, but particularly combinations of those two, especially with *be, be able to*, and/or humans, as in (56).

(56) *If it weren't for Mr. Fox, I wouldn't have been able to go to college.*

This counterfactual conditional construction also dislikes a variety of more specific uses of benefactives, e.g. those with *die* (57) or those with positive polarity, especially with *have*, *be*, and *be able to*, as in (58).

(57) *If it weren't for this new method, my parents wouldn't have died. They were suffering a lot; I am glad this method worked.*
(58) *If it weren't for them, I would have had a lot of distractions.*

3.3 *If It Wasn't For Constructions*

3.3.1 Preferred Co-occurrences

This construction has some very general preferences: one for malefactives in general, one for humans in general, and one for negative polarity in general, but there are also a variety of smaller subtypes arising from combinations of these, such as:

– negative polarity with *die* and malefactives (59):

(59) *If it wasn't for Mike, she wouldn't have died. I really hate Mike!*

– negative polarity with *be*, *have*, and *know*, especially with humans and benefactives (60):

(60) *If it wasn't for Bill, Aaron wouldn't have a lot of money.*

– humans with *die* and malefactives (61):

(61) *If it wasn't for Dr. Lewis, she wouldn't have died.*

In addition, *if it wasn't for* also likes positive polarity with *be*, especially with humans or benefactives, as in (62):

(62) *If it wasn't for Mr. Smith, I would be a burglar.*

3.3.2 Dispreferred Co-occurrences

This construction also has very general dispreferences: benefactives in general, positive polarity in general, and inanimates, as well as combinations of these in particular with *be able to*, as in (63):

(63) *If it wasn't for the new wall, I would be able to see you. I am glad I won't see you anymore, dude!*

4 Discussion

The results presented in Section 3 allow for several observations regarding *if not for-if it weren't/wasn't for* constructions in American English.[5] In this section, we discuss these results and consider their linguistic implications.

4.1 *Linguistic Implications*

As discussed in Section 2, the semantics of benefactive and malefactive *if not for* vs. *if it weren't/wasn't for* constructions are different. While the former means 'in the absence of X, Y would have failed to accomplish Z or would not have avoided Z', the latter means 'in the absence of X, Y would not have failed to accomplish Z or would have avoided Z.' Accordingly, we expected that benefactive *if not for* and *if it weren't/wasn't for* constructions would (prefer to) occur with protases with humans, animates, or inanimates when the apodoses

5 In response to one reviewer's suggestion, we actually also performed one analysis in which the *if it wasn't for* and the *if it weren't for* constructions were conflated as one construction, but the results for the *if not for* construction were interpretationally identical to our analysis and the results for the two conflated constructions combined features from the two separate ones, which means, if anything, the conflation lost some power to discriminate between *if it wasn't for* and *if it weren't for*.

appear with detriment verbs (e.g. *kill, die*) and show positive polarity. Moreover, we expected that they would (prefer to) occur with protases with humans, animates, or inanimates when the apodoses appear with accomplishment verbs (e.g. *have, get, become*) and show negative polarity. On the other hand, we expected that malefactive *if not for* and *if it weren't/wasn't for* constructions would (prefer to) occur with protases with humans when the apodoses appear with accomplishment verbs (e.g. *have, get, become*), and show positive polarity. Finally, we expected that malefactive *if not for* and *if it weren't/wasn't for* constructions would prefer protases with humans when the apodoses appear with state verbs (e.g. *be*) and would show negative polarity. In what follows, we explore whether these predictions hold.

As for *if not for* constructions, Table 3 provides a summary of the types and antitypes attested in the present study. Note that an empty cell is to be interpreted as 'abstracting away over all levels that cell could contain.' For example, the last row means '*if not for* disprefers malefactives with inanimates and positive polarity *no matter the apodosis lemma*.' Also, in the discussion below we use a curly-brackets set notation to represent clusters of features that, in Table 3, are rows (see also Figure 1).

A closer look reveals that our expectations hold for benefactives. First, the animacy of the protasis NP may be human, animate, or inanimate when the apodosis appears with detriment verbs, such as *lose*, *hurt*, *fail*, *kill*, and *die*, and shows positive polarity, as in (64).

(64) *If not for Jacob, I would have died. Honestly, I was afraid of dying. I am not ready to see God and my family*.

This cluster correlates with the meaning of this sort of benefactive *if not for* construction. Semantically, it should be understood as: 'in the absence of X, Y would not have avoided Z'. Put another way, in this scenario, an undesirable situation Z (i.e. a situation causing harm or damage to Y) was averted by the support/help/presence of the benefactor X, which can show any sort of animacy. These constructions are similar to avertive 'lest' constructions. Accordingly, we call them *avertive/precautioning constructions* (see Lichtenberk, 1995: 298). In this construction, the purpose of an action consists in preventing an undesirable situation from happening.

Another verb lemma occurring in benefactives with an avertive/precautioning function is that of *seem* (see Table 3). Intriguingly, this verb does not qualify as a detriment verb. The semantics of this verb is that of 'to give the effect of being; to be judged to be'. This verb has been classified as a deductive epistemic judgement verb in that it is used for indicating the type of evidence

TABLE 3 Summary of types and antitypes of *if not for* constructions

CONSTRUCTION	(Dis) Pref	CX TYPE	NP TYPE	APODOSIS LEMMA	POLARITY
If not for	pref	benefactive		*lose, hurt*	positive
If not for	pref	benefactive	human	*fail, kill*	positive
If not for	pref	benefactive	inanimate	*seem*	positive
If not for	pref	benefactive	human	*die*	positive
If not for	pref	malefactive	human	*be*	negative
If not for	pref		human	*be able to*	negative
If not for	pref	benefactive	human or inanimate	*exist*	negative
If not for	pref	benefactive	inanimate	*know*	negative
If not for	dispref	malefactive	human or inanimate	*have*	negative
If not for	dispref	benefactive		*die*	negative
If not for	dispref		human	*die*	negative
If not for	dispref		human	*be able to, know*	positive
If not for	dispref	benefactive		*be able to, know*	positive
If not for	dispref	malefactive		*know*	positive
If not for	dispref	benefactive		*have*	positive
If not for	dispref	malefactive	inanimate		positive

that speakers have to say that 'X seems like Y' (Palmer, 2001: 24). The question is: why does *seem* occur here if its semantics does not harmonize with the semantics of benefactives involving an avertive/precautioning function? The examples in our corpus indicate that *seem* tends to be followed by similative 'like' phrases appearing with adjectives involving a detriment, as in (65) and (66), where an undesirable appearance was averted. Under this analysis, the semantics of *seem* aligns with the semantics of benefactive constructions involving an avertive/precautioning function.

(65) *If not for the medicine that my mother gave you, you'd seem like a dead guy*.

(66) *If not for your phone, you'd seem like a lost dog*.

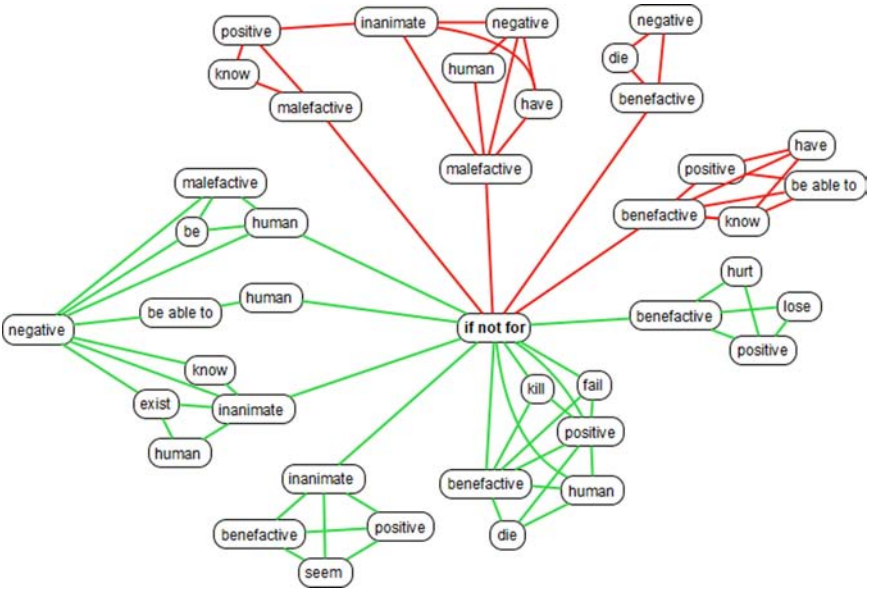

FIGURE 1 A visual representation of Table 3 (green and red lines connecting features of types and antitypes respectively)

Second, unlike benefactive *if not for* constructions with an avertive/precautioning function, this type of counterfactual construction can also show an achievement function. As can be seen in Table 3, this is attested with apodoses with verbs, such as *exist* and *know*, and with negative polarity apodoses, as in (67) and (68) respectively. Accordingly, this harmonizes with our expected meaning of this construction: 'in the absence of X, Y would have failed to accomplish Z.' For lack of a better term, we call these constructions *achievement constructions*.

(67) *If not for Charles, our company wouldn't exist in this competitive world.*

(68) *If not for this video, I wouldn't know how to behave properly now that I am the president of the firm.*

As can be seen in Table 3, only malefactive *if not for* constructions meaning: 'in the absence of X, Y would have avoided Z', are among the types. Note that unlike avertive/precautioning benefactives, in this type of *if not for* construction an undesirable situation was not averted. For lack of a better term, we call these constructions *anti-precautioning/avertive constructions*. As is shown in

Table 3, this construction prefers protases with humans, apodoses with *be*, and negative polarity apodoses.

As for *be*, there are 32 cases where *be* is followed by a prepositional phrase (PP) denoting an undesirable location as (69).

(69) *If not for you, I wouldn't be in this cold place. I really hate this!*

In this scenario, an undesirable situation Z could have been avoided if the protasis participant X had not intervened. This indicates that this participant must be a maleficiary human entity given that it represents a volitional causation that affects another entity adversely. Put another way, the affectee is affected negatively by a maleficiary human entity in a direct fashion. Accordingly, this cluster harmonizes with our expected meaning of this construction.

There are a number of *if not for* construction antitypes that deserve closer attention. First, as can be seen in Table 3, *if not for* constructions do not like to appear with the cluster {benefactive, *be able to/know/have*, and positive polarity}. As was shown above, benefactive *if not for* constructions with positive polarity have a precautioning/avertive function and strongly prefer detriment verb lemmas. The semantics of *be able to/know/have* does not align with the precautioning/avertive function given that they are more related to an accomplishment. Accordingly, this may explain why these verbs are dispreferred in benefactive *if not for* constructions with positive polarity.

Second, as discussed above, benefactive *if not for* constructions occurring with *die* prefer positive polarity apodoses. There, they denote a precautioning/avertive function. Without positive polarity apodoses, it may be difficult for speakers to come up with a construction in which benefaction and detriment verbs are involved, as in (70) (repeated for convenience). This may explain why the cluster {benefactive, *die*, negative polarity} is dispreferred.

(70) *If not for Dr. Morales, my mother wouldn't have died. We are very grateful because Dr. Morales found out a way to help her see God. After years of fighting cancer, she can now rest in peace.*

Third, regardless of the polarity of the apodosis, we expected that malefactives will occur with humans in the protasis because they represent maleficiaries that affect someone Y negatively in a direct fashion. Accordingly, it seems reasonable to propose that this is why the cluster {malefactive, inanimate, and positive polarity} is dispreferred.

We now turn to types and antitypes of *if it weren't for* constructions. For types, it is clear from Table 4/Figure 2, there are fewer clusters than those attested for *if not for* constructions.

Moreover, the preferred types of this counterfactual construction involve only malefactives and nearly all dispreferred types are benefactives according to the statistical metrics we used and described in Section 3. Two observations are in order here. First, *if it weren't for* constructions prefer the cluster {malefactive, *know*, positive polarity}, as shown in Table 4/Figure 2. Accordingly, these constructions should be understood as: 'in the absence of X, Y would

TABLE 4 Summary of types and antitypes of *if it weren't for* constructions

(Dis) Pref	Construction	Cx Type	NP Type	Apodosis Lemma	Polarity
pref	*If it weren't for*	malefactive			positive
pref	*If it weren't for*	malefactive		*know*	positive
pref	*If it weren't for*	malefactive	human		positive
dispref	*If it weren't for*				negative
dispref	*If it weren't for*	benefactive			
dispref	*If it weren't for*	benefactive	human	*be*	negative
dispref	*If it weren't for*	benefactive		*die*	
dispref	*If it weren't for*	benefactive		*have, be, be able to*	positive

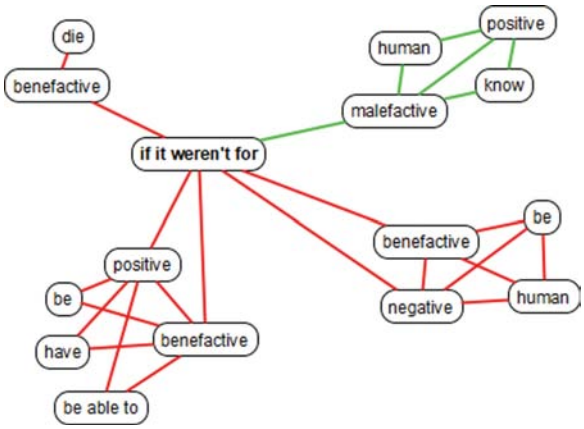

FIGURE 2 A visual representation of Table 4

not have failed to accomplish Z.' Note that unlike benefactive constructions with an achievement function (e.g. *if not for him, I wouldn't have become president*), in this type of construction a desirable situation was not accomplished, as in (71). For lack of a better term, we call these constructions *nonachievement constructions*. Second, the fact that malefactive *if it weren't for* constructions prefer humans in their apodoses in general is not surprising given that, as was mentioned above, they represent maleficiaries that affect some Y negatively in a direct fashion.

(71) *If it weren't for you, I would have known how to become a good father. I can't believe you didn't help me!*

As can be seen in Table 4, most clusters of *if it weren't* antitypes disprefer benefactives. Two antitypes deserve careful attention. First, *if it weren't for* constructions do not like the cluster {benefactive, human, *be*, negative polarity}, as in (72). At the current stage of our work, it is not clear to us why this cluster is dispreferred by *if it weren't for* constructions. It is worth noting that in the example in (72), *be* should be considered an accomplishment verb. Accordingly, it harmonizes with our expected meaning of benefactive *if not for-if it weren't/wasn't for* with negative polarity apodoses.

(72) *If it weren't for Mike, I wouldn't be a successful person.*

Second, for benefactive *if not for* constructions with an avertive/precautioning function, we have shown that they prefer detriment verbs, such as *die*, *kill*, and *hurt*, etc. This may explain why *have* and *be* are dispreferred in benefactive *if it weren't for* construction with positive polarity apodoses (see Table 4). However, note that *have* and *be* may show a detriment function, as in the examples in (73) and (74). Accordingly, the question is: why are *have* and *be* dispreferred in this scenario if they can be used for expressing a detriment function? The reason why *have* and *be* are dispreferred in constructions like (73) and (74) may have to do with conceptualization/construal. This indicates that speakers prefer to conceptualize the situations expressed in the construction in (73) as: *if it weren't for him, I would have hurt my leg*, and the situations expressed in the construction in (74) as: *if it weren't for Simon, I'd have died.* Semantic conventions emerge "from recurrent conceptualizations of the same or similar experiences that become associated with particular lexemes and constructions" (Diessel, 2019: 29).

(73) *If it weren't for him, I would have had a hurt leg*.
(74) *If it weren't for Simon, I'd have been a dead guy*.

To round off the discussion, let's discuss *if it wasn't for* types and antitypes. The multivariate extension of collostructional analysis adopted here shows one type that clearly supports our initial expectations, i.e. *if it wasn't for*, benefactive, humans, *be*, *have*, and *know*, and negative polarity apodoses, as in (75) (see Table 5). The fact that *if it wasn't for* prefer this cluster is not surprising: *be*, *have*, and *know* can be considered accomplishment verbs and harmonize with our expected meaning of this construction: 'in the absence of X, Y would have failed to accomplish Z.'

(75) *If it wasn't for Jackson, I wouldn't be the writer I am today*.

Two types, at first glance, do not seem to support our initial expectations. First, *be* is clustered with benefactive *if it wasn't for* constructions showing positive polarity apodosis. Recall that we expected that this type of construction will prefer detriment verbs (e.g. *die*, *hurt*) given that it has an avertive/precautioning function. Accordingly, the question is: why does *be* occur here if its

TABLE 5 Summary of types and antitypes of *if it wasn't for* constructions

(Dis) Pref	Construction	Cx Type	NP Type	Apodosis Lemma	Polarity
pref	*If it wasn't for*	malefactive			
pref	*If it wasn't for*		human		
pref	*If it wasn't for*				negative
pref	*If it wasn't for*	malefactive		*die*	negative
pref	*If it wasn't for*	benefactive	human	*be, have, know*	negative
pref	*If it wasn't for*	malefactive	human	*die*	
pref	*If it wasn't for*		human	*be*	positive
pref	*If it wasn't for*	benefactive		*be*	positive
dispref	*If it wasn't for*	benefactive			
dispref	*If it wasn't for*				positive
dispref	*If it wasn't for*		inanimate		
dispref	*If it wasn't for*	benefactive	inanimate	*be able to*	positive

semantics does not harmonize with that of benefactives involving an avertive/precautioning function? In most examples in our dataset, *be* is followed by an adjective denoting loss, damage, harm, and injury, as in (76).

(76) *If it wasn't for you, I would be dead.*

Under this analysis, the semantics of *be* aligns with that of benefactive constructions involving an avertive/precautioning function. As argued above, conceptualization is one of the main driving forces behind the construction of meaning so speakers may prefer to express detriment with *be* in benefactive constructions involving an avertive/precautioning function. As Croft (1991: 108–109) explains, "we organize our experience, encoded by the semantics of lexical items, in a particular way in order to communicate it to our interlocutors." Second, *if it wasn't for* constructions also cluster with {malefactive, *die*, and negative polarity}, as shown in Table 5. Based on our initial expectations, we would not expect *die* to appear in malefactive *if it wasn't for* constructions with negative polarity apodoses. This stems from the fact that this cluster prefers verbs that harmonize with the anti-precautioning/avertive function and not with detriment verbs that are preferred by constructions with a precautioning/avertive function (e.g. *if not for him, I would have died*).

Finally, mention should be made of the following *if it wasn't for* antitype. As can be seen in Table 5/Figure 3, *if it wasn't for* constructions do not like the

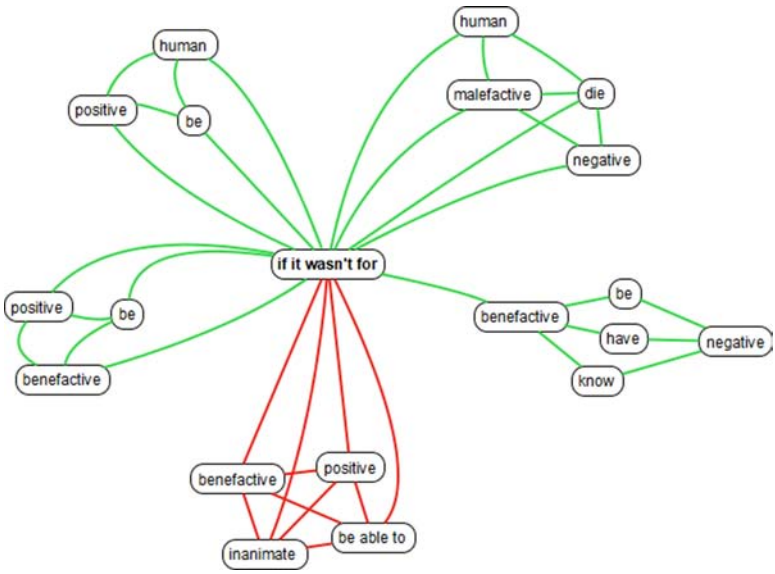

FIGURE 3 A visual representation of Table 5

cluster {benefactive, inanimate, *be able to*, and positive polarity} as in (77). As has been shown in this section, the cluster {benefactive, positive polarity}, prefers detriment verbs given that its meaning is similar to that of: 'in the absence of X, Y would not have avoided Z.' The lemma *be able to* is not a detriment verb used to express an undesirable situation that was averted. Accordingly, it seems reasonable to propose that this is why *be able to* is dispreferred in benefactives with a precautioning/avertive function.

(77) *If it wasn't for the new wall, I would be able to see you. I am glad I won't see you anymore, dude!*

5 Concluding Remarks

From a usage-based Construction Grammar perspective, a lexeme can be, and typically is, combined with a construction if it fits the semantic specifications of a particular slot in a construction, a notion often referred to as the semantic coherence principle (Goldberg, 1995: 50). On a most basic level, our work is further evidence that the associations between individual lexemes and slots of constructions one would expect from this principle do exist and that they are semantically motivated. However, while a traditional multiple distinctive collexeme analysis would have restricted itself to one slot in each construction, we used a now truly multivariate extension of collostructional analysis, which indicates that lexical choices are also influenced by other linguistic variables or, more broadly and ambitiously speaking, that we should begin to move on from the traditional kind of collostructional analysis towards a more powerful and versatile approach that can take many more slots (e.g. not just ApodosisLemma but also NPType) and much other contextual/functional information of one or more constructions' usage into consideration at the same time (also CxType and Polarity).

We exemplified our proposal and its potential with an exploration of one counterfactual conditional construction traditionally neglected in the literature: *if not for-if it weren't/wasn't for* counterfactual constructions. All instances of the three constructions were annotated for their semantics (e.g. whether the construction indicates a benefactive or malefactive meaning), the animacy of the NP in the protasis and the verb lemma and the polarity (positive vs. negative) of the apodosis and submitted to our new method, which returned a variety of interesting clusters, or configurations.

In summary, *if not for* constructions with an avertive/precautioning function were significantly attracted to the cluster {benefactive, noun phrases with any sort of animacy, detriment verbs (e.g. *die, kill, hurt*), positive polarity}. We argued that this cluster harmonizes with our expected meaning of this benefactive *if not for* construction, semantically understood as: 'in the absence of X, Y would not have avoided Z.' *If not for* constructions with an achievement function represented another type. We showed that this type preferred the cluster {accomplishment verbs (e.g. *exist*, *know*), negative polarity}, which harmonizes with our expected meaning of this construction: 'in the absence of X, Y would have failed to accomplish Z.' For malefactive *if not for* constructions with an anti-precautioning/avertive function, we showed that they are significantly attracted to the cluster {human, *be*, negative polarity}, which aligns with our expected meaning of this construction: 'in the absence of X, Y would have been able to avoid an unpleasant situation.'

As for *if it weren't for* constructions, the preferred types of this counterfactual construction involved only malefactives. In particular, we showed that *if it weren't for* constructions with a nonachievement function prefer the cluster {malefactives, *know*, positive polarity}, which harmonizes with our expected meaning of this construction: 'in the absence of X, Y would not have failed to accomplish Z.'

For *if it wasn't for*, benefactives having an achievement function were significantly attracted to the cluster {human, *be*, *have*, and *know*, negative polarity}, which aligns with our expected meaning of this construction: 'in the absence of X, Y would have failed to accomplish Z.'

The cluster types discussed above should be taken as evidence that speakers seem to have entrenched various meso-constructions that are characterized by cross-clausal associations (e.g. NPTYPE, APODOSISLEMMA) and other contextual/functional information (POLARITY). From a usage-based Construction Grammar perspective, these mental representations with varying degrees of schematicity are stored in taxonomic networks. It is worth noting that one might expect *if it wasn't for* to be more similar to *if it weren't for* than to *if not for*, and *if it weren't for* to be more similar to *if it wasn't for* than to *if not for*.

6 It is worth mentioning that, in response to one reviewer's comment with regard to what a *vanilla collexeme analysis* might show, we did do a traditional multiple collexeme analysis. The results of that analysis led to interpretable results only for the *if not for* construction – the other two constructions' top verbs exhibited no interpretable patterns. The constructions studied here, which are not argument structure constructions with very clear semantics (like, e.g. the ditransitive) differ less in terms of their lexical preferences than their other associated characteristics, which is why the vanilla analysis does not offer more or more helpful generalizations than the more fine-grained analysis we proposed.

However, an initial and as yet exploratory comparison of the different cluster types discussed in this paper suggests that this is not the case. The only type that partially overlaps with another type is that in which achievement *if not for* constructions show the cluster {*exist*, *know*, negative polarity}, and achievement *if it wasn't for* constructions show the cluster {human, *be*, *have*, and *know*, negative polarity}.

The multivariate collostructional analysis adopted here also enabled us to identify a number of antitypes of *if not for-if it weren't/wasn't for* counterfactual constructions, which we also interpreted with regard a variety of semantic constraints.[6]

In spite of the richness of the results, there are aspects relevant to the analysis of *if not for-if it weren't/wasn't for* counterfactual constructions that the present research could not address. Accordingly, they remain to be investigated by future studies, and in what follows some potentially fruitful areas are mentioned. First, one variable that was not taken into account in the present study, but seems to be relevant to the analysis of *if not for-if it weren't/wasn't for*, is the order of the protasis and apodosis. Protases tend to precede apodoses because the protasis refers to a situation that is conceptually prior to the one expressed in the apodosis (Diessel, 2008: 469). In our study, there are cases in which apodoses do not follow an iconic order. Interestingly, a closer look reveals that they may be characterized as apodosis with a pragmatic function, i.e. a clause that refers to linguistic or non-linguistic aspects of the ongoing social interaction/speech situation itself. Formally, this is reflected by non-declarative clause types, i.e. interrogative clause types (e.g. *what would you have done if not for her?*). Pragmatically, this yields a framing potential that is relevant to the organization of ongoing social interaction (Hampe & Gries, 2018). It remains to be analyzed how this variable interacts with the types and antitypes attested in the present study.

Second, our study did not utilize any kind of predictive modeling to explore the degree to which the three constructions exhibit distributional patterns that are so distinctive that they are in fact predictive and, more specifically, whether the *if it wasn't for* and the *if it weren't for* constructions are allostructions. We commented on this above on the basis of our interpretation of the findings, but this could also be tested more quantitatively, e.g. by using predictions based on predictive modeling (as per Hampe & Gries, 2018) or on association rules (see Jensen & Gries, under review).

We hope that the method adopted here will be valuable to other usage-based construction grammarians to uncover how cross-clausal dependencies in biclausal constructions also interact with other linguistic variables, which opens up new avenues of research for the future.

References

Byrne, Ruth M. J. 2002. Mental models and counterfactual thoughts about what might have been. *Trends in Cognitive Sciences* 6: 426–431.

Croft, William. 1991. *Syntactic categories and grammatical relations: The cognitive organization of information*. Chicago: Chicago University Press.

Croft, William and Alan Cruse. 2004. *Cognitive linguistics*. Cambridge: Cambridge University Press.

Dahl, Östen. 2008. Animacy and egophoricity: Grammar, ontology and phylogeny. *Lingua* 118: 141–150.

Davies, Mark. 2008. *The Corpus of Contemporary American English* (COCA). Available online at https://www.english-corpora.org/coca/.

Declerck, Renaat and Susan Reed. 2001. *Conditionals. A comprehensive empirical analysis*. Berlin/New York: Mouton de Gruyter.

Diessel, Holger. 2008. Iconicity of sequence. A corpus-based analysis of the positioning of temporal adverbial clauses in English. *Cognitive Linguistics* 19: 457–482.

Diessel, Holger. 2019. *The grammar network. How linguistic structure is shaped by language use*. Cambridge: Cambridge University Press.

Divjak, Dagmar S. 2019. *Frequency in language: Memory, attention and learning*. Oxford: Oxford University Press.

Fillmore, Charles J. and Paul Kay. 1999. *Construction grammar*. Berkeley: University of California.

Goldberg, Adele E. 1995. *Constructions: A construction grammar approach to argument structure*. Chicago: University of Chicago Press.

Gries, Stefan Th. 2009. *Statistics for linguistics with R: A practical introduction*. Berlin/New York: Mouton de Gruyter.

Gries, Stefan Th. and Anatol Stefanowitsch. 2004a. Extending collostructional analysis: A corpus-based perspective on 'alternations'. *International Journal of Corpus Linguistics* 9: 97–129.

Gries, Stefan Th. and Anatol Stefanowitsch. 2004b. Co-varying collexemes in the into-causative. In M. Achard and S. Kemmer (eds.), *Language, culture, and mind*, 225–236. Stanford: Center for the Study of Language and Information.

Gries, Stefan Th. and Anatol Stefanowitsch. 2010. Cluster analysis and the identification of collexeme classes. In S. Rice and J. Newman (eds.), *Empirical and experimental methods in cognitive/functional research*, 73–90. Stanford: Center for the Study of Language and Information.

Hampe, Beate and Stefan Th. Gries. 2018. Syntax from and for discourse II: More on complex-sentences as meso-constructions. In B. Hampe and S. Flach (eds.), *Yearbook of the German Cognitive Linguistics Association*, 115–142. Boston/Berlin: De Gruyter Mouton.

Hoffmann, Thomas, Jakob Horsch, and Thomas Brunner. 2019. The more data, the better: A usage-based account of the English comparative correlative construction. *Cognitive Linguistics* 30: 1–36.

Jensen, Kim E. and Stefan Th. Gries. under review. Multivariate collostructional analysis via association rules.

Jing-Schmidt, Zhuo. 2017. What are they good for? A constructionist account of counterfactuals in ordinary Chinese. *Journal of Pragmatics* 113: 30–52.

Kittilä, Seppo and Fernando Zúñiga. 2010. Introduction: Benefaction and malefaction from a cross-linguistic perspective. In F. Zúñiga and S. Kittilä (eds.), *Benefactives and malefactives typological perspectives and case studies*, 1–28. Amsterdam/Philadelphia: John Benjamins.

Klomps, Koen. 2021. *A corpus study on verb tenses in Dutch and English Conditionals*. Ph.D. dissertation, Utrecht University.

Lichtenberk, Frantisek. 1995. Apprehensional Epistemics. In J. Bybee and S. Fleischman (eds.), *Modality in grammar and discourse*, 293–327. Amsterdam/Philadelphia: John Benjamins.

Louw, Bill. 1993. Irony in the text or insincerity in the writer? The diagnostic potential of semantic prosodies. In M. Baker, G. Francis, and E. Tognini-Bonelli (eds.), *Text and technology*, 157–176. Amsterdam/Philadelphia: John Benjamins.

Morley, John and Alan Partington. 2009. A few frequently asked questions about semantic – or evaluative – prosody. *International Journal of Corpus Linguistics* 14: 139–158.

Newman, John. 2011. Corpora and cognitive linguistics. *Brazilian Journal of Applied Linguistics* 11: 521–559.

Nordquist, Dawn. 2004. Comparing elicited data and corpora. In M. Achard and S. Kemmer (eds.), *Language, culture, and mind*, 211–223. Stanford: Center for the Study of Language and Information.

Olguín Martínez, Jesús and Nicholas Lester. 2021. A quantitative analysis of counterfactual conditionals in the world's languages. *Italian Journal of Linguistics* 33: 147–182.

Palmer, F. R. 2001. *Mood and modality* (2nd edition). Cambridge: Cambridge University Press.

Silverstein, Michael. 1976. Hierarchy of features and ergativity. In R.M.W. Dixon (ed.), *Grammatical categories in Australian languages*, 112–171. Canberra: Australian Institute of Aboriginal Studies.

Smith, Aaron and Dawn Nordquist. 2012. A critical and historical investigation into semantic prosody. *Journal of Historical Pragmatics* 13: 291–312.

Stefanowitsch, Anatol and Stefan Th. Gries. 2003. Collostructions: Investigating the interaction between words and constructions. *International Journal of Corpus Linguistics* 8: 209–243.

Stefanowitsch, Anatol and Stefan Th. Gries. 2005. Covarying collexemes. *Corpus Linguistics and Linguistic Theory* 1: 1–43.

Stefanowitsch, Anatol and Susanne Flach. 2020. *Too big to fail* but *big enough to pay for their mistakes*: A collostructional analysis of the patterns [too ADJ to V] and [ADJ enough to V]. In G. Corpas Pastor and J. P. Colson (eds.), *Computational phraseology*, 247–272. Amsterdam/Philadelphia: John Benjamins.

Van Linden, An and Jean-Christophe Verstraete. 2008. The nature and origin of counterfactuality in simple clauses: Cross-linguistic evidence. *Journal of Pragmatics* 40: 1865–1895.