Constructional schemas in variation: modelling contrastive negation

Running head: Constructional schemas in variation

Olli O. Silvennoinen
University of Helsinki
Department of Languages
P.O. box 24 (Unioninkatu 40)
00014 University of Helsinki, Finland

olli.silvennoinen@helsinki.fi

Abstract

This paper discusses constructional variation in the domain of contrastive negation in English, using data from the British National Corpus. Contrastive negation refers to constructs with two parts, one negative and the other affirmative, such that the affirmative offers an alternative to the negative in the frame in question (e.g. *shaken, not stirred*; *not once but twice*; *I don't like it—I love it*). The paper utilises multiple correspondence analysis to explore the degree of synonymy among the various constructional schemas of contrastive negation, finding that different schemas are associated with different semantic, pragmatic and extralinguistic contexts but also that certain schemas do not differ from each other in a significant way.

Keywords

contrastive negation; corpus linguistics; focus; grammatical variation; multiple correspondence analysis; Principle of No Synonymy; usage-based linguistics.

1. Introduction

In cognitive and construction-based linguistics, it is generally held that "if two ways of saying something differ in their words or their arrangement they will also differ in meaning" (Bolinger 1977:1). Known as the Principle of No Synonymy (Goldberg 1995:67), this tenet follows from the idea that even abstract grammatical constructions are meaningful and provide subtly different construals of events and situations. Thus according to Goldberg (1995:33), the English ditransitive construction implies a successful metaphorical transfer (e.g., *Mary taught Bill French* implies that *Bill learned French*), while the double object construction is neutral as to whether the transfer actually takes place (e.g., *Mary taught French to Bill* does not entail that *Bill learned French*). On the constructional account, argument structure constructions like this have rich semantics that are independent of the verbs because they have an experiential basis (Lakoff 1987): transfer, for example, is a "humanly relevant scene", and the ditransitive construction is the linguistic emblem of

this scene. Furthermore, Goldberg predicts that if two formally distinct constructions are semantically synonymous, they must be pragmatically non-synonymous, and vice versa, pragmatics being defined as "particulars of information structure, including topic and focus, and additionally stylistic aspects of the construction such as register" (Goldberg 1995:67).

There are sounds of discord. Even after an extensive statistical analysis, Granvik and Taimitarha (2014) find that certain topic-marking prepositions in Swedish are semantically and stylistically in free variation. Similarly, Cappelle (2006) and Perek (2015) have found that certain constructional alternations (e.g., the English particle alternation: *She turned off the TV* / *She turned the TV off*) can be accounted for as more general-level constructions that have two possible instantiations. This raises the question whether the strict trade-off between semantic and pragmatic synonymy is warranted. In particular, we may ask whether the Principle of No Synonymy holds for constructions that are semantically very schematic such as constructions whose main function is related to information structure.

This paper considers the problem of near-synonymous constructions by focusing on one schematic construction family, which I shall call contrastive negation, following McCawley (1991; see also Gates Jr. & Seright 1967). I define contrastive negation as combinations of a negative part with a parallel affirmative part that offers an alternative to the negated one in the frame in question. The six major constructional schemas for contrastive negation in English are presented in (1)–(6):

(1) [*not X but Y*][1]
Not stirred but shaken.

(2) [*not X, Y*]
Not stirred, shaken.

(3) Expanded [*not X, Y*]
The drink is not stirred – it is shaken.

[1] The word *not* is not the only word possible in this slot; other negators are also possible. It is used in naming the constructional schemas because of convenience.

(4) [*X not Y*]
Shaken, not stirred.

(5) [*X and not Y*]
Shaken and not stirred.

(6) Expanded [*X not Y*]
The drink is shaken. It is not stirred.

What separates contrastive negation from previously studied grammatical alternations is the large number of constructions that are formally analogous to one another. Schemas (1)–(3) place the negated element first, schemas (4)–(6) place it second. Schemas (2) and (4) are asyndetic, i.e., they have no overt coordinator, while also being sub-clausal, whereas (1) and (5) are syndetic and generally sub-clausal.[2] (3) and (6) are also asyndetic but clausal; for them, I have retained McCawley's (1991) term "expanded".

On the other hand, these six constructional schemas cross-cut three semantic types, as defined on the basis of overtly coded scalarity (Dik et al. 1981:59–68).[3] The basic type is replacive, which has no overtly coded markers of scalarity in either conjunct. The second type is additive, in which there is a scalar expression (e.g., *just, only, simply, merely*) attached to the negated conjunct, and the third is restrictive, in which a scalarity marker attaches to the affirmed conjunct. Examples (7)–(8) illustrate the [*not X but Y*] schema with replacive, additive and restrictive semantics, respectively:

(7) Savoy is *not an archaeologist but an explorer*. [AAT, 189] [4]

(8) 'There is a moral crisis in sport, *not only in Canada but on a worldwide scale*,' he said. [A2S, 281]

[2] The [*not X but Y*] schema is only prototypically sub-clausal, while [*X and not Y*] always operates on the constituent level.

[3] The three semantic types are based on the three types of corrective focus in Dik and colleagues' overall typology of focus phenomena, though they call them "replacing", "restricting" and "expanding". Dik et al. also treat example (i) as "selective focus" which is a non-corrective type of focus, "because A's presupposition […] is not incorrect" (Dik et al. 1981:62). This distinction will not have consequences for the analysis presented here, and therefore the issue is left aside in the present paper.

(i) A: Did John buy coffee or rice?
B: John bought coffee, not rice. (Dik et al. 1981:62, modified)

[4] The three-character code in square brackets refers to a file in the BNC and the number to the s-unit (i.e., orthographic sentence or turn) in that file. Examples without a code or a source reference are constructed by the author.

(9) It is particularly illogical that this kind of argument should be coming from politicians who, in other contexts, would be the first to argue, and rightly, that Vietnam is *not some kind of monster State, but merely a ramshackle and inefficient one that has lost its way*. [AA1, 13]

All six constructional schemas can be used for all three semantic types, and in fact all combinations are attested. This alone is equal to eighteen potential constructions listed in an individual speaker's constructicon, i.e., list of constructions. This makes the quantitative treatment of contrastive negation challenging, and from a psycholinguistic point of view it is also possible that not all of those constructions are entrenched in a speaker/writer's constructicon.[5]

The literature contains very little information on contrastive negation from the point of view of constructional variation in any language. The existing descriptions of the phenomenon are mainly based on anecdotal, introspective or experimental data, and they offer few clues as to why a specific construction might be chosen in a given context. The studies that do consider usage data have only done so regarding individual constructions (e.g., Toosarvandani 2010). For this reason, this study takes an exploratory approach to the constructional variation of contrastive negation. Using data from the British National Corpus (BNC), I model the behaviour of the six constructional schemas with multiple correspondence analysis (MCA). The research questions that I ask are the following:

- What usage patterns are associated with which constructional schemas?
- How close are the schemas to one another in terms of their usage patterns?
- Are there true synonyms among the schemas?

My working assumption is that the schemas exemplified in (1) through (6) are surface generalisations (Goldberg 2006:23), i.e., abstractions over more specific constructions. However, whether these generalisations are actually represented in speakers' constructicons is another matter. I adopt Traugott and Trousdale's (2013:16) conception of levels of constructional schematicity, in which constructions range from highly abstract constructional schemas (macro-constructions) to

[5] This, of course, does not mean that speakers without a given construction would not be able to recognise it when faced with a relevant construct in discourse. Fillmore et al. (1988:502) make a distinction between knowing the meaning of a construction and figuring it out. For instance, an expanded negative-second construction of the additive type would readily be understood as such even if the hearer did not possess this particular construction as an entrenched item in their constructicon.

mid-level sub-schemas (meso-constructions) and lexically specific fixed expressions (micro-constructions). I consider the construction types (e.g., the [*X not Y*] schema) to be constructional schemas, and the specific combinations of a construction type with a semantic type (e.g., the additive [*X not only Y*] schema, as in *shaken, not only stirred*) to be potential sub-schemas – potential because an individual may not store all such combinations in their constructicon. Speakers also have the possibility of recording micro-constructions in the domain of contrastive negation (e.g., *shaken, not stirred* is probably listed in most English speakers' mental constructicons). In general, lower-level constructions are more psychologically real to speakers and using a lower-level construction does not necessarily mean that the higher-level schema is evoked (see Hilpert 2015; Perek 2015). Low-level schemas also have richer semantics than the constructional schemas, being less abstracted from individual constructs used in discourse. The present study will explore which schemas typically have the status of construction in an average speaker's constructicon – and which do not. Therefore, I shall mostly use the term "schema", only using the term "construction" when I deem there to be sufficient reasons for positing a conventionalised unit.[6]

The structure of the paper is as follows. In section 2, I present the data and methods. Section 3 then discusses the variables used in the analysis. Section 4 presents the results and section 5 concludes the paper.

2. Data and methods

Given the diversity in the expression of contrastive negation, corpus data is needed to examine the variation between the schemas. My study will use the British National Corpus (BNC) as data. The BNC is a multi-genre corpus of roughly 100 million words of British English collected mostly in the early 1990s. Around 10% of it is spoken, the rest written. Specifically, I shall be using the national broadsheet component of the corpus. The broadsheet data amounts to approximately 3

[6] This should not be taken as denial of the fact that speakers often store redundant patterns (Langacker 1987:29). Given that these storage patterns are a matter of individuals, they fall beyond the scope of this paper, which is based on population-level corpus data.

million words (for more on the dataset and how it was collected, see Silvennoinen (2017)). The corpus was accessed through the BNCweb interface[7] (for documentation, see Berglund et al. 2002; Hoffmann et al. 2008). Approximately 2,000 tokens of contrastive negation were collected by first searching for negators and then manually finding instances of contrastive negation among all other negative construction types.[8] Of these, I extracted those cases in which the affirmative and negative parts differ by one syntactic element (see McCawley 1991:189). This means that cases such as (10)–(13) were excluded from the analysis. Example (10) is an additive contrast of two states of affairs ('the Xhosas were out in force' and 'a convention of the entire Mass Democratic Movement had been called'); however, these two states of affairs are structured differently. A similar situation obtains in (11). In example (12), the contrast is between two ways of distributing participants to participant roles. Example (13) contains two versions of a similar contrast in which the same process ('winning/losing an election') is profiled differently. In all these cases, the contrasted elements could not be stated as alternative ways of filling a single syntactic position and therefore these kinds of cases were left out.

(10) Yesterday not only were the Xhosas out in force, it seemed that a convention of the entire Mass Democratic Movement – essentially the exiled ANC's internal representatives – had been called. [A1G, 54]

(11) In fact, I only met one decorator … a young couple wanted to buy a piece and I offered them a discount as they hadn't got a decorator – at which she drew herself up and said: '*We are not married. I am his interior decorator.*' [A5J, 94–95]

(12) It is *the Americans* who remake *foreign* films, *not the other way round*. [A2J, 323]

(13) AN OLD maxim of British politics states that *oppositions do not win elections, governments lose them*. After Thursday, a new maxim ought to state that *governments do not win elections, oppositions lose them*. [AK2, 549–550]

There were several reasons for restricting the dataset in this way. Firstly and most importantly, this restricts the focus to cases in which all or most of the six schemas are possible. For example, (10) is

[7] See http://corpora.lancs.ac.uk/BNCweb/.

[8] Not all of those 2,000 tokens can be categorised into the six constructional schemas in (1)–(6). The entrenched constructional schemas notwithstanding, the grammar of contrastive negation is to some extent open-ended, a feature that will not be explored in this paper (but see Silvennoinen (2017) for discussion).

difficult if not impossible to express using the [*X not Y*] schema. This reduces the potential for knockout constraints, i.e., combinations of variables in which a given schema does not appear at all (Paolillo 2002:30). Secondly, this conforms to previous studies of contrastive negation (Gates Jr. & Seright 1967; McCawley 1991), which have defined it syntactically. On one hand, this means that some of the variability in the domain of contrastive negation is left out of the present study. On the other, the study targets those aspects of variation that are most likely to apply to all six schemas that are being investigated. This procedure yielded 1,831 cases of contrastive negation, which are the dataset proper of this study. Table 1 shows the distribution of the six schemas in the dataset.

@@ Insert Table 1 here

Table 1. Raw frequencies in the data

	N
[*not X but Y*]	813
[*not X, Y*]	80
Expanded [*not X, Y*]	459
[*X not Y*]	351
[*X and not Y*]	99
Expanded [*X not Y*]	29

The main statistical method used in this study is multiple correspondence analysis (MCA), a form of correspondence analysis (CA). MCA, like all forms of CA, is an exploratory technique for analysing patterns among discrete variables (Greenacre 2017; Glynn 2014; Le Roux & Rouanet 2010; chapter 19 in Levshina 2015). In other words, it "helps one find which usage-features [i.e., levels of variables] co-occur with other usage-features, giving a map of their overall patterning" (Glynn 2014:444). The results of MCA are typically presented as a biplot, i.e., a two-dimensional space into which the variable levels and the observations are plotted. In this respect, it resembles

Principal Components Analysis, which is used for continuous variables. For the purposes of the analysis, the dataset was coded for a number of variables, which will be presented in section 3. Each variable was either discrete or discretised for the purposes of the analysis: in other words, they all consist of two or more categories or levels in statistical parlance. The dimensions produced in MCA typically do not correspond to any individual variable but are agglomerates of several variable levels, which are represented in the dimension at varying degrees. The closer two observations are in the biplot, the more similar their levels. If we assume that similarity in usage indicates similarity in function, we would expect synonymous constructions to pattern similarly and non-synonymous constructions to pattern apart from one another in the biplot.

The input to CA is in many respects similar to that of logistic regression, which has established itself as a common methodology in the study of lexical and constructional (near-)synonymy (see, e.g., Arppe 2008; Gries 2003). In addition, CA is in many respects similar to clustering techniques (see, e.g., Baayen 2008:118–148), which are arguably more established in statistics. Both are exploratory techniques that visualise multivariate datasets. However, Glynn (2014:478) argues that there are two advantages to CA. First, clusters are presented as discrete while CA provides a visualisation that is more appropriate for a phenomenon as gradient as synonymy. Second, since variable levels can also be shown in the biplot, CA offers a more transparent view into the factors that lead to certain constructions being judged more similar than others. From a practical point of view, CA is more suitable for the present study in which the schemas differ greatly in terms of frequency and are not very numerous, thus producing clusters that might be uninformative (because of the relative poverty of the set of schemas) or even misleading (because of data scarcity).

3. The variables

This section discusses the variables that were included in the analysis. Some of the variables are associated with a directional hypothesis ("this level is positively/negatively associated with these schemas") while others are not. The latter are included because of the exploratory nature of the

study – since there have not been prior studies on contrastive negation from a variationist point of view, we do not yet know which set of features will yield results. However, the selection should not be arbitrary or too broad as MCA is blind to whether a correlation is spurious or not (Glynn 2014:444–445). Therefore, care has been taken to select variables that are conceptually independent of one another. The mnemonic names of the variables appear in small capitals.

3.1. Variables related to meaning and form

3.1.1. Semantic type

Semantic type (SEMTYPE) is an independent variable with three levels: replacive, additive and restrictive, as exemplified in (7)–(9). The division into semantic types was made on the basis of overt marking of scalarity: if scalarity was marked on the negative part, the construct was coded as additive. If it was marked on the affirmative part, the construct was coded as restrictive. If it was marked on neither, the construct was coded as replacive. As a result, even cases that display lexicalised scalarity were coded as replacive if the scalarity was not marked in the form of the construct.[9] Consider (14) in which the contrasted elements are an entrenched pairing on the scale of number of times: {*once* < *twice*}.

(14) Architects/Engineers RMJM have been flying high recently, rubbing shoulders with royalty, *not once but twice* within a week. [A24, 146]

This is a replacive construct but the contrasted elements are such that an additive construction (*not only once but twice*) would also be possible.

3.1.2. Target of negation

Probably most of the previous research on contrastive negation concerns metalinguistic negation (Anscombre & Ducrot 1977; Horn 1985, 1989; Carston 1996). In Horn's seminal account, metalinguistic negation is the marked alternative of descriptive negation. While descriptive negation targets the propositional content, metalinguistic negation targets its expression, including

[9] If the contrasted items formed an implicational scale, they were coded as negation targeting non-propositional content (see section 3.1.2), irrespective of scalarity marking. Therefore, the present approach has the advantage of making semantic type and the target of negation conceptually independent of one another, the former being based on the form of the construct and the latter on its interpretation.

implicatures and presuppositions that the expression might ordinarily be expected to have. I follow McCawley (1991) by considering that contrastive negation may be both descriptive and metalinguistic, and metalinguistic negation may be both contrastive and non-contrastive.[10]

Geurts (1998) has proposed a more fine-grained classification of "the mechanisms of denial", which is presented in (15)–(18).

(15) Proposition denials:
If Ramon hadn't been Spanish but French, he would still beat his donkey. (Geurts 1998:280)

(16) Presupposition denials:
Barney didn't take his WIFE to Acapulco—he isn't even married—but his GIRLfriend. (Geurts 1998:287)

(17) Implicature denials:
A: Julius had six beers.
B: He didn't have six beers: he had at least seven. (Geurts 1998:287)

(18) Form denials:
A: Kurt swallowed a whole to[mɑ:]to.
B: He didn't swallow a to[mɑ:]to but a to[meɪ]to. (Geurts 1998:287)

Proposition denials correspond to Horn's descriptive negation, while presupposition, implicature and form denials together correspond to metalinguistic negation. Presupposition denials deny a proposition; in (16), for instance, the presupposition that Barney is married, which would ordinarily survive negation, is denied. Implicature denials deny an implicature, particularly a scalar one as in (17). Literally speaking, if one has had seven beers, one has also had six since the latter is a subset of the former. However, saying that Julius has had six beers implicates that this is the maximal value that the speaker is entitled to say. Since B thinks that Julius has had "at least seven", they produce a denial of the scalar implicature. Finally, form denials deny the appropriateness of the linguistic form that has been used; in (18), what is objected to is the pronunciation.[11]

[10] Indeed, arguably the most famous example of metalinguistic negation – *The king of France is not bald, because there is no king of France* (Horn 1985:121) – is not contrastive.

[11] In cases of presupposition and implicature denial and sometimes even form denial, the negation actually targets both propositional content and its expression. Thus, the distinction between descriptive and metalinguistic negation is not a

Additive constructs are often used in situations similar to implicature denials, as in (19):

(19) He didn't only have six beers; he had at least seven.

Cases like this are classified together with (16) since they too concern the denial of a scalar implicature. However, these cases are not strictly speaking metalinguistic because owing to the adverb *only*, (19) is literally true.

Metalinguistic negation has been shown to affect the choice of negative-contrastive construction in languages with a negator that is specialised for contrastive negation. In Malay/Indonesian, for instance, descriptive cases of contrastive negation may be formed either with the specialised *bukan* or the standard negator *tidak*, but if the negation is metalinguistic, *bukan* must be used (Kroeger 2014). A similar pattern has been observed in some varieties of Arabic (Mughazy 2003; Chatar-Moumni 2008) and in Mandarin (Yeh 1995). Thus it makes sense to test whether metalinguistic negation or a similar category is a factor in construction choice in English. The variable is termed the target of negation (NEGTARGET). Its levels are called proposition denial and non-proposition denial, and the latter category also includes cases that are strictly speaking instances of descriptive negation, like (19), because of their functional equivalence with implicature-denying metalinguistic negation. By doing this, I keep SEMTYPE and NEGTARGET conceptually independent of one another.[12]

3.1.3. Negator

Contrastive negation also shows variation in the negator. The paradigm of English negators is quite extensive, but for present purposes I decided to follow Tottie's (1991) two-way distinction into *not*-negation (*not* and *n't*) and *no*-negation (*no* and all other "absolute negators" (Huddleston & Pullum 2002:788): *nobody*, *no one*/*no-one*, *nothing*, *none*, *neither*, *nor*, *never*, *nowhere*). The reason for pooling all negators besides *not* together is that contrastive negation is strongly associated with *not*-

clear-cut one between content on one side and expression on the other. As Carston (1996, 1998) points out, a more apt formulation might be that metalinguistic negation is echoic: the negated content is mentioned rather than used.

[12] Presupposition, implicature and form denials have been pooled together to avoid data scarcity.

negation (Tottie 1991:161, 163–170; Silvennoinen 2017) and therefore the other negators make only a limited number of appearances in the dataset. However, this does not preclude the possibility that some schemas might be more tolerant of *no*-negation. The variable negator (NEG) is thus included in the study, and it has two levels.

3.2. Variables related to information structure

3.2.1. Weight

Ordering of the negated and asserted elements is one of the most conspicuous features differentiating negative-contrastive constructions. A prime candidate to determine ordering in any construction family is the weight of the elements, usually operationalised in terms of their length. Weight and length are obviously also related to form but here they are grouped as an information-structural variable because they concern the way the contrast is packaged in the construction. In English, length has been found to be operative in the choice of comparison strategy, the genitive alternation and particle placement, among others (Gries 2003:83–85, 110; Szmrecsanyi 2006:97–98, 101). Length taps into the end-weight principle (Behaghel 1909/1910), which in turn is thought to facilitate processing (Hawkins 2003) or production (Wasow 1997). Given the strong evidence from other areas of grammar, we may hypothesise that when the negated element is longer than the asserted one, it tends to follow it, and vice versa.

Weight was considered as the length of the focus of affirmation minus the length of the focus of negation. It was operationalised in terms of the number of words, since this has been found to be a good enough measure of syntactic complexity, in addition to being the easiest of such measures to utilise (Szmrecsanyi 2004). Example (20) illustrates the analysis:

(20) The suffering of God is not FOC[eternal and infinite]; it is FOC[human and limited and the same kind of suffering as that of Auschwitz or of cerebral meningitis]. [A3F, 46]

In (20), the focus of affirmation (*human and limited and the same kind of suffering as that of Auschwitz or of cerebral meningitis*) is 17 words and the focus of negation (*eternal and infinite*) 3 words. The length difference is thus 17–3=14 words.

The definition of focus adopted here conflates two notions: focus as defined in the study of information structure and focus as defined in the study of negation. Among the notoriously many definitions of focus in information structure, I shall follow Lambrecht (1994:207), who defines focus as "the element of information whereby the presupposition and the assertion DIFFER from each other" (emphasis in the original). Here, presupposition is taken to mean those propositions that "the hearer already knows or is ready to take for granted at the time the sentence is uttered" while assertion is defined as the propositions that "the hearer is expected to know or take for granted as a result of hearing the sentence uttered" (Lambrecht 1994:52). In (20), for instance, the presupposition is that the suffering of God exists and is the topic of discussion, while the assertion is a characterisation of it. The focus, then, is how the suffering of God is characterised, both negatively and affirmatively.

I shall assume that the focus of clausal negation coincides with Lambrecht's definition of focus. However, in cases of constituent negation, Lambrecht's definitions need to be broadened. I analyse the non-expanded forms of contrastive negation as paired focus constructions that are on par with the *let alone* construction, for instance (Fillmore, Kay & O'Connor 1988). That is, both the asserted and the negated element are focal even in contrastive negation that is effected through constituent negation.[13]

[13] Recall the difference between the scope and focus of negation (Blanco & Moldovan 2013; Huddleston & Pullum 2002:790–799; Quirk et al. 1985:787–794). Consider (i):

(i) Win or lose, I guarantee SCO[we will **not** return FOC[with our tails between our legs]], but FOC[with a hell of a lot of dignity and applause from you blokes].

In (i), the negation is clausal even though the construction is non-expanded. McCawley (1991) calls such constructions "anchored". What is central here is that scope and focus do not coincide. The whole clause is in the scope of the negation, i.e., any clause element could potentially be the focus of negation (Quirk et al. 1985:789–790). However, the

Because of differences among the construction types, determining the length of the foci of negation and assertion was not always straightforward. The following rules were followed. First, the negators and focusing adverbs were not counted among words. However, when the negator had a function in addition to negating and was part of the contrasted clause element, it was counted, the logic being that a corresponding affirmative word would have been counted as part of the asserted focus. For instance, in (21), the negator *not* and the restrictive adverb *only* are not counted as part of the focus of negation, while in (22), the negative conjunction *nor* is included in the word count because the corresponding affirmative (*or*) would receive the same treatment.

(21) Television, both BBC's coverage for 20 years and ITV's for the past five, not only FOC[attracted lucrative sponsorships] but FOC[developed a greater interest among the young]. [A2S, 120]

(22) What Westminster MPs tend not to understand is that the beneficiaries of this process have not been FOC[their cousins in the Strasbourg Parliament, nor even the European Commission], but FOC[the Council of Ministers]. [AA1, 526]

Second, with regard to cases in which the contrasted elements were finite verb phrases, I excluded the auxiliary *do*, which is part of the Standard Negation Construction in the simple tenses in English. This was to maximise equivalence between the negated and asserted elements. Consider (23), in which *doesn't* is not counted as being part of the focus:

(23) 'I'm voting for John Major because he is a great Prime Minister. Because he doesn't FOC[punish success], he FOC[promotes it],' she purred. [AJD, 505–506]

Third, elements that were outside of the clause structure proper were not counted. This criterion picked out orthographically marked parentheticals, for instance, as in (24):

(24) With me it has done well not only FOC[in a sheltered, sunny position] (the best conditions) but also FOC[in a place where it has only a few hours of direct sunshine each day]. [AK6, 1541]

Since MCA requires all variables to be discrete, the numeric length difference values were binned into three groups: those in which the focus of affirmation was longer, those in which the focus of

contrastive construction disambiguates the focus by providing a parallel alternative. As to contrastive negation created with constituent negation, I shall assume that the scope and focus of negation are equivalent.

negation was heavier, and those in which the foci were of equal length. The variable WEIGHT thus has three levels: aff-heavy, neg-heavy and balanced.

3.2.2. Focus structure

In addition to weight, foci can be analysed in terms of how they are realised syntactically. Lambrecht (1994) determines three possible focus structures in a clause. The first type is the predicate-focus structure, exemplified in (25) and analysed in (26):[14]

(25) A: What happened to your car?
B: My car/It broke DOWN. (Lambrecht 1994:223)

(26) Predicate-focus structure (Lambrecht 1994:226)

Presupposition:	'speaker's car is a topic for comment x'
Assertion:	'x = broke down'
Focus:	'broke down'
Focus domain:	VP

The predicate-focus structure is the unmarked choice among focus constructions. It is a topic—comment structure in which the VP provides the focal information.

The second type is the narrow-focus structure, exemplified in (27) and analysed in (28):[15]

(27) A: I heard your motorcycle broke down?
B: My CAR broke down. (Lambrecht 1994:223)

(28) Narrow-focus structure (Lambrecht 1994; Van Valin 2005)

Presupposition:	'speaker's x broke down'
Assertion:	'x = car'
Focus:	'car'
Focus domain:	NP

The narrow-focus structure picks an argument as the focal element. It is compatible with correction, as in (27), and by extension contrastive negation, although this is not a necessary context of occurrence for the construction.

---

[14] In examples (25), (27) and (29), small capitals indicate sentence stress.

[15] Lambrecht (1994) calls this type "argument-focus structure". However, as Van Valin (2005) points out, this type also includes adjuncts. I have also classified sub-phrasal foci as narrow-focus structures:

(ii) An overwhelming majority were also in favour of the mixed economy, suggesting that the reformers are opposed not FOC[because of their policies], but FOC[because of their image]. [A2X, 479]

(iii) Negotiations which exclude the legitimate leaders of our people, which the government appears to be contemplating, will FOC[intensify], not FOC[resolve], our country's crisis. [A3D, 438]

In (ii), the *because*-phrases are adjuncts, and in (iii), the foci are non-finite parts of a VP.

The third type is the sentence-focus structure, exemplified in (29) and analysed in (30):

(29) A: What happened?
B: My CAR broke down. (Lambrecht 1994:223)

(30) Sentence-focus structure (Lambrecht 1994:233)

Presupposition:	—
Assertion:	'speaker's car broke down'
Focus:	'speaker's car broke down'
Focus domain:	S

In the sentence-focus structure, the whole sentence is focal and there is no presuppositional background to which the structure is anchored. Both narrow-focus and sentence-focus structures are marked compared to the predicate-focus structure, which would explain why their forms are identical in (27) and (29). Since sentence-focus is by definition not a property of any individual constituent of a clause but must always concern several parts of it, cases with sentence-focus fall outside of the scope of the present study. Focus structure (FOCUSSTR) is therefore a variable with two levels: predicate-focus and narrow-focus.

The question of which focus structures are permissible with which schemas has been the topic of some disagreement in the literature. Some focus structures seem impossible with certain schemas. A case in point is the [*X not Y*] schema combined with finite verb phrases, i.e., predicate-focus structures. The *Cambridge Grammar of the English Language* judges this combination to be acceptable in (31). Gates and Seright provide the opposite judgement in (32), in which they claim that only an Expanded schema with overt subjects is acceptable in case of a negative-second construct with predicate-focus. Such differences of opinion suggest that the [*X not Y*] schema is at least statistically dispreferred with predicate-focus, and thus I hypothesise that it is not associated with it in the corpus data.[16]

[16] It may be wondered why I approach this as a question of focus structure rather than phrase type, since Gates and Seright as well as Huddleston and Pullum frame their discussions around the concept of finite VPs. Consider (iv):
(iv) I'm not crying for myself; I'm crying for the others. [A7Y, 373]

(31) The night turned viciously cold under a sky crowded with stars that *shone, not twinkled*, in the diamond-clean air. (cf. *shone, didn't twinkle) (Huddleston & Pullum 2002:812, emphasis added)

(32) a. He ran—he did not walk—to the station.
b. *He ran, not walked, to the station. (Gates Jr. & Seright 1967:140)

By contrast, some of the other schemas studied in this paper readily and uncontroversially combine with the predicate-focus structure; see, e.g., (23) above.

3.2.3. Activation

The activation of discourse elements is generally taken to influence construction choice. Negation as a whole has been argued to be contingent on the activation or at least accommodation of the negated proposition (Givón 1978; Tottie 1991:33–38). Since activation refers to the mental states of the speaker/writer and the hearer/reader and therefore cannot be observed directly, the literature includes various kinds of operationalisations. In their seminal study on the dative alternation, Bresnan et al. (2007) treat what they term "accessibility" of the theme and the recipient as a binary variable (given/not given), although they do not elaborate on their coding criteria. Their results show that nongivenness of recipient and nongivenness of theme help predict the choice between the prepositional dative and double object constructions. Szmrecsanyi (2006:94, 135–136) also treats information structure (in his words, "information status" or "news value") as a binary variable. In his study on five morphosyntactic alternations, information structure is evoked in two: genitive choice and particle placement. In the former, the possessor and the possessum were coded as given if the head lemmas were mentioned in the last 100 words, while in the latter the givenness of the direct object was determined on the basis of whether its referent had been mentioned in the previous five sentences. According to Szmrecsanyi's analysis, news value predicts particle placement but not genitive choice, though in the latter case it may be that other variables cover the same effect (Szmrecsanyi 2006:96–97, 102–104, 141–142). Interestingly, it is the cruder metric of previous five

---

The literature has no consensus on how to determine whether the focus should be only the NP within the PP (Lambrecht 1994) or the whole PP (Blanco & Moldovan 2013). A blanket policy either way risks biasing the data, but anything else would be highly subjective and therefore unreliable. The present approach avoids this problem.

sentences that produces the better result in Szmrecsanyi's study. A similar result was achieved in Gries's study on particle placement, which operationalises the "news value" of the direct object as both a categorical binary (mentioned in any of the previous discourse/not mentioned) and an ordinal variable (distance to the last mention in terms of the number of intervening sentences; number of past mentions) (Gries 2003:50, 73). Gries also addresses the "relevance" of the referent by tracking the number of its subsequent mentions and the distance to the next mention, although these are far less effective as predictors (Gries 2003:92–95).

Prince (1981) notes that is it very difficult if not impossible to say whether referents in written discourse are brand-new to the context or whether they are contextually inferable on the basis of the topic of the text, for instance. For this reason, the present study follows the previous multivariate research by treating activation as a binary variable, based on whether the referents (or designata in the case of elements that are not strictly speaking referential, such as adjectives; see Lambrecht (1994:37)) of the contrastive elements have been mentioned in the previous four sentences. In addition, deictic and anaphoric references were counted as references to activated concepts. The variable is labelled NEGACTIVATED, and it has two levels: activated (labelled as "active") and not activated (labelled as "non-active"). The analysis is illustrated in (33), in which the negated concept (*morality*) is mentioned in adjectival form (*moral*) three sentences earlier:

(33) But it would be hard to argue, say, that Greeks are less moral because of the appalling corruption of and greed demonstrated by senior figures in the government and party of the socialist Andreas Papandreou. Dr Runcie acknowledged that 'wealth creation' was the necessary precondition for 'doing all the things we ought to do' but added that there was no automatic connection between wealth creation and a happy society. Amen to that. Market economics is about *efficiency, not morality*. [A1F, 107–110]

The window of four sentences was chosen as a matter of convenience: it is the largest amount of co-text that can be exported from the BNCweb interface. Thus, in this study activation refers to immediate prior mention (or co-presence in the case of deictic references). Naturally, it is only an indirect means of tapping into activation, but it is in line with previous corpus-based research.

Activated concepts are generally expressed earlier rather than later in a construction, and for this reason my hypothesis is that activated concepts favour constructions in which the negated element precedes the asserted one. In practice, this means that the negative-second schemas should have fewer activated concepts as the negated element than the others.

3.3. Contextual variable: genre

Finally, the various genres included in the dataset are included in the analysis. Contrastive negation appears to be a genre-sensitive feature of language (see Silvennoinen 2017): argumentative genres such as editorials favour it while sports reports have very low levels of it. The realisation of negation in English overall is sensitive to the difference between speech and writing (Tottie 1991). The variable GENRE has eight levels: arts, commerce, editorial, miscellaneous, report, science, social and sport. The division into genres is part of the corpus metadata. It is based on Lee (2001).

3.4. Summary

The variables are summarised in Table 2. It includes the codes of the variables (in small capitals), the codes of the variable levels (in typewriter font) and the hypotheses related to each variable, if any. The codes will figure in the exposition of the statistical analysis below.

@@ Insert Table 2 here

Table 2. Variables, variable levels and hypotheses

Type of variable	Variable	Levels	Hypotheses
Variables related to meaning and form	SEMTYPE (semantic type)	`rep` (replacive) `add` (additive) `rst` (restrictive)	–
	NEGTYPE (type of negation)	`prop` (negation that targets propositional content) `non_prop` (negation that targets non-propositional content instead of or in addition to propositional content)	–
	NEG (negator)	`not` (*not*-negation)	–

		`no` (*no*-negation)	
Variables related to information structure	WEIGHT (length difference between the foci of affirmation and negation)	`aff-heavy` (focus of affirmation longer than focus of negation) `balanced` (the foci of affirmation and negation are of equal length) `neg-heavy` (focus of negation longer than focus of affirmation)	`aff-heavy` is associated with negative-first schemas `neg-heavy` is associated with negative-second schemas
	FOCUSSTR (focus structure)	`narrow` (narrow-focus) `predicate` (predicate-focus)	`pred` is not associated with the [*X not Y*] schema
	NEGACTIVATED (activation of the focus of negation)	`active` (activated) `non-active` (not activated)	`a` is associated with negative-first schemas
Contextual variable	GENRE	`arts` (arts/cultural material) `commerce` (commerce & finance) `editorial` (personal & institutional editorials & letters-to-the-editor) `misc` (miscellaneous material) `report` (home & foreign news reportage) `science` (science material) `social` (material on lifestyle, leisure, belief & thought) `sports` (sports material)	–

The table shows that there are both purely exploratory variables and variables that also test a specific hypothesis or hypotheses. The latter are all related to information structure variables, which is natural since these are based on previous studies on other grammatical domains. The variables related to meaning and form as well as the contextual variable GENRE have effects that are specific to contrastive negation and thus harder to predict.

4. Results

4.1. Univariate and bivariate results

Before the actual multivariate analysis, let us look at the variables individually, i.e., the univariate results of the analysis, and pairwise, i.e., the bivariate results. All statistical analyses were

conducted in the open-source statistical programming environment R (R Core Team 2016). Space does not permit an exhaustive treatment of all the variables and their interactions, so we shall simply consider the variable of interest, CONSTRUCTIONTYPE in relation to the other variables. Table 3 presents an overview of the distributions of the variables in the constructional schemas.

@@ Insert Table 3 here

Table 3. Overview of variables

		[*not X but Y*]	[*not X, Y*]	Expanded [*not X, Y*]	[*X not Y*]	[*X and not Y*]	Expanded [*X not Y*]
SEMTYPE	rep	409	15	320	313	81	27
	add	378	–	75	24	9	–
	rst	26	65	64	14	9	2
NEGTARGET	prop	692	77	411	324	93	29
	non-prop	121	3	48	27	6	–
NEG	not	782	50	406	342	94	28
	no	31	30	53	9	5	1
WEIGHT	aff-heavy	553	51	276	83	27	15
	balanced	129	19	58	164	31	7
	neg-heavy	131	10	125	104	41	7
FOCUSSTR	narrow	683	78	254	351	98	15
	predicate	130	2	205	–	1	14
NEGACTIVATED	active	168	20	91	84	10	5
	non-active	645	60	368	267	89	24
GENRE	arts	154	20	76	51	7	2
	commerce	72	7	38	29	10	7
	editorial	76	4	33	28	13	3
	misc	244	32	158	128	36	11
	report	154	6	88	78	21	3
	science	19	1	12	9	–	2
	social	33	2	29	15	5	–
	sports	61	8	25	13	7	1

MCA detects patterns of associations among variables in an exploratory way. Unlike in some other statistical methods such as linear regression, collinearity between the variables is not in itself a

problem for the model. However, since MCA is blind to whether the associations are statistically significant or not, we should determine statistical significance independently (Weller & Romney 1990:57). I did this using the chi square test. The results of the tests are reported in Appendix 1. Most of the tests came out with a significant result, with only around one result in three being not significant. Crucially, each variable has associations with other variables, and all variables except NEGACTIVATED are associated with CONSTRUCTIONTYPE. This suggests that it is appropriate to perform MCA on the dataset, as the correlations between CONSTRUCTIONTYPE and the other variables are unlikely to be down to chance (see Glynn 2014:444).

The variables related to meaning and form are all exploratory as they are not associated with a hypothesis. For SEMTYPE, replacives are the most frequent alternative for all schemas except for [*not X, Y*], which prefers restrictive semantics. Also, for [*not X but Y*] the difference between replacives and additives is slight. These findings are broadly in line with the fact that replacives are formally and functionally the unmarked option. As to NEGTARGET, proposition denial as the unmarked option is more prevalent than the other types of denial put together. [*not X but Y*] has the highest degree of non-proposition denial. When it comes to NEG, *not*-negation is unsurprisingly the more common option for all schemas.

The variables related to information structure come with specific hypotheses, which are largely borne out by the data. WEIGHT seems to show a preference for end-weight for most of the constructions, especially for [*not X but Y*]; [*X not Y*], however, seems to favour contrasted elements that are balanced in terms of weight. Of the two types of FOCUSSTR, narrow-focus is more common for all schemas, although for Expanded [*X not Y*] the two options are virtually tied. Three schemas ([*X not Y*], [*X and not Y*] and, perhaps surprisingly, [*not X, Y*]) display a dramatic dispreference for predicate-focus. For the negative-second schemas, this is expected because of the disagreement of previous studies regarding such constructs (see section 3.2.2). As for NEGACTIVATED, all schemas prefer non-activated foci of negation. This variable is not significant when cross-tabulated with

CONSTRUCTIONTYPE, and therefore, activation does not appear to predict construction choice very strongly in the domain of contrastive negation, contrary to the dative alternation, for instance.

With regard to the only extralinguistic variable considered here, GENRE, there are few trends that would appear particularly noteworthy, although the variable does have a significant association with CONSTRUCTIONTYPE.[17]

4.2. Multiple correspondence analysis

We can now move on to the multivariate statistical analysis of the data. MCA was performed using the package `FactoMineR` (Lê, Josse & Husson 2008). Supplementary analysis was done using the `ca` package (Nenadic & Greenacre 2007). The analysis proceeded in several steps. First, the statistical analysis was performed; for this, we get a numerical output of the dimensions created in the analysis. The dimensions are created automatically on the basis of co-occurrence patterns in the data. Second, the numerical output was interpreted, i.e., the dimensions were related to the underlying variables. Third, the analysis was visualised using biplots. This will help in the interpretation of the analysis. Lastly, the model was validated using logistic regression.

The first step was to create the numerical analysis. MCA summarises the patterns found among the variables of the dataset by converting them into distances along a number of dimensions. The dimensions are described in Table 4. The dimensions differ in terms of how much of the variance they explain. They are ordered so that dimension 1 explains the largest share of the variance, dimension 2 the second largest, and so on. The measure of the amount of variance explained by a dimension is called principal inertias in CA, and they are given in the column labelled "Value" in Table 4. The table then reports two percentages of the variance explained: "%" is the amount of variance explained according to regular MCA, "Adjusted %" the amount of variance explained according to adjusted MCA, which is a sub-form of MCA that provides a realistic assessment of the

[17] The findings on GENRE have been discussed more fully in Silvennoinen (2017).

contributions (Levshina 2015:382; Glynn 2014:450; Greenacre 2017).[18] On the basis of the latter, the cumulative amount of variance explained is given in the next column. Finally, the scree plot in the final column shows graphically how much each of the dimensions contributes to the model. The number of dimensions that are to be considered in subsequent steps of the analysis is based on where the proportion of explained variance elbows. After the third dimension, there is virtually no increase in the proportion of explained variance. Therefore, in what follows I shall only consider dimensions 1, 2 and 3.

@@ Insert Table 4 here

Table 4. Principal inertias

Dimension	Value	%	Adjusted %	Cumulative adjusted %	Scree plot of the adjusted %
1	0.005202	9.552	33.2	33.2	**********
2	0.004301	9.290	27.4	60.6	**********
3	0.000715	7.737	4.6	65.2.	**
4	0.000170	7.188	1.1	66.3	
5	1.2e-050	6.807	0.1	66.4	
6	1e-06000	6.710	0.0	66.4	
7	1e-06000	6.705	0.0	66.4.	

As said, each dimension is an aggregate of several variables (or, to be more precise, their levels). Table 5 shows how much each variable contributes towards the first three dimensions.

@@ Insert Table 5 here

[18] In non-adjusted MCA, the proportion of explained variance is underestimated because the algorithm relates each level of each variable not only to each level of every other variable but also to all the remaining levels of the same variable. Since the levels are mutually exclusive (e.g., the SEMTYPE can only be `rep`, `rst` or `add`, not a combination of these), this results in a lot of zeroes in the underlying data matrix. These zeroes bring the explained variance down. Adjusted MCA, also called joint MCA, solves this problem. See chapter 19 in Greenacre (2017) for discussion.

Table 5. Contributions of the variables to the dimensions

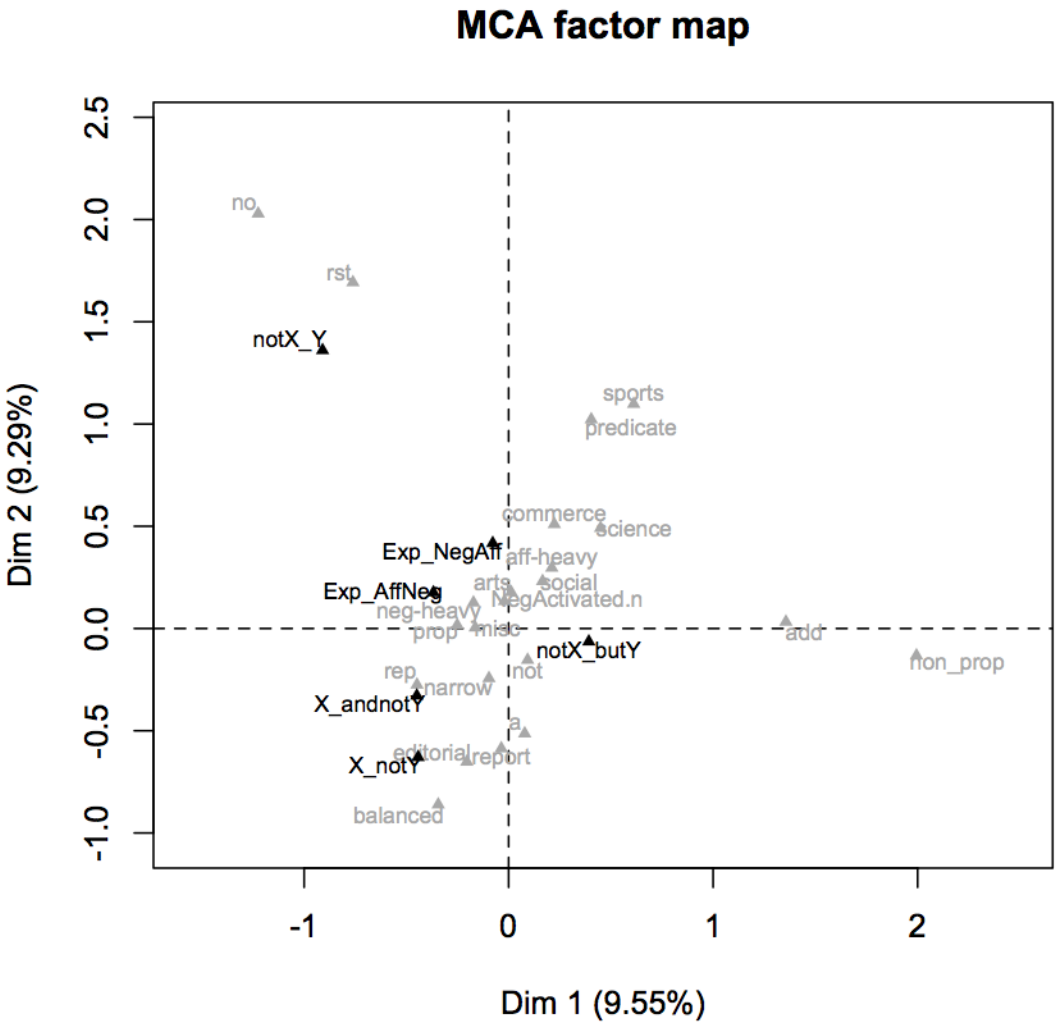
	Dimension 1	**Dimension 2**	**Dimension 3**
NEG	0.114	0.312	0.067
SEMTYPE	0.673	0.330	0.120
WEIGHT	0.058	0.217	0.303
NEGTARGET	0.502	0.002	0.019
NEGACTIVATED	0.002	0.069	0.244
FOCUSSTR	0.039	0.249	0.096
GENRE	0.047	0.215	0.313

I shall now go through the three dimensions and give them interpretations by relating them to the kinds of cases that each dimension differentiates from one another. The contributions of the variables as a whole are given in Table 5, while information on individual levels can be found in the figures.

Figure 1 is a biplot of dimensions 1 and 2. The figure shows both the positions of the schemas (in black) and of the levels of the variables (in grey). The schemas appear to form four groups: the Expanded schemas are one, the phrasal negative-second schemas ([*X not Y*] and [*X and not Y*]) are another, and [*not X, Y*] as well as [*not X but Y*] are both one. However, the expanded schemas and the phrasal negative-second schemas are quite close to one another, and they appear to form a gradient with the Expanded [*not X, Y*] schema at one end and [*X not Y*] at the other. (In this biplot and all subsequent ones, the Expanded [*not X, Y*] and [*X not Y*] schemas are labelled "Exp_NegAff" and "Exp_AffNeg" for the sake of convenience.)

@@ Insert Figure 1 here

Figure 1. Biplot of dimensions 1 and 2



As shown in the variable contributions listed in Table 5, dimension 1 is mostly based on SEMTYPE and NEGTARGET, in a descending order of importance. As to SEMTYPE, additives have a strongly positive and restrictives a strongly negative association with dimension 1; replacives are also slightly negatively associated with it. Non-proposition denial is associated positively with it, while proposition denial has an equally strong negative association. Of the schemas, [*not X but Y*] is strongly associated with the positive pole of dimension 1, while [*not X, Y*] has a notable negative association with it. In sum, the prototype of an example with a positive association to dimension 1 is an additive [*not X but Y*] construct which exhibits metalinguistic negation or another type of

scalar implicature denial. This is not a random cluster of features: the additive [*not X but Y*] construction is often used for denying an upper-bound scalar implicature, and these cases are classified as non-proposition denial. In (34), for instance, *Le Monde* is characterised as an institution, which is more notable than being a mere newspaper:

(34) Le Monde is not just a newspaper but an institution. [A2D, 94]

The prototype case for the negative pole of dimension 1 is a restrictive, proposition-denying instance of the [*not X, Y*] schema. This is line with the observation that restrictive constructs are seldom metalinguistic in my data. One such case is (35):

(35) There was *no debate on the national executive, merely an expression of prejudice and fear of change*. [A30, 219]

Dimension 2 has a more even spread of contributing variables, as shown in Table 5. The variables that contribute the most towards dimension 2 are SEMTYPE, NEG, FOCUSSTR, WEIGHT and GENRE, in a descending order of importance. SEMTYPE is treated somewhat differently by dimensions 1 and 2: whereas dimension 1 pitted additives against both replacives and restrictives, dimension 2 puts restrictives at the positive end and replacives at the negative one. As to NEG and FOCUSSTR, *no*-negation and predicate-focus are positively and *not*-negation and narrow-focus negatively associated with the dimension. Interestingly, contrasted elements with both kinds of unbalanced weights (affirmative heavier than negative, negative heavier than affirmative) are associated with the positive end of dimension 2 while balanced contrasts are associated with the negative end. Of the genres, sports is positively associated with the dimension while editorial and report have negative scores with dimension 2. The [*not X, Y*] schema has a positive association with the dimension, while [*X and not Y*] and especially [*X not Y*] are negatively associated with it. The positive prototype of dimension 2 is thus similar to the negative prototype of dimension 1. A prototype of a negative value on dimension 2 has a replacive [*X not Y*] construct with narrow-focus and *not*-negation that appears in an argumentative or reportative text and in which the contrasted

elements are equally long. This is a natural cluster since this schema is strongly attracted to *not*-negation and is virtually impossible with predicate-focus. This prototype is exemplified in (36):

(36) 'The priority of MAFF must become FOC[*food*], *not* FOC[*farming*],' the policy review says. [A1J, 524]

I now move on to dimension 3. Figure 2 shows the biplot for dimensions 1 and 3, and Figure 3 the biplot for dimensions 2 and 3. These plots are not as easy to interpret as the first one. In both figures, the most dramatic result concerns [*not X, Y*], which is the far north-east of Figure 2 and the far north-west of Figure 3. All other schemas appear quite close to origo in both plots, which indicates that their values in dimension 3 are quite moderate.

@@ Insert Figure 2 here

@@ Insert Figure 3 here

Figure 2. Biplot of dimensions 1 and 3

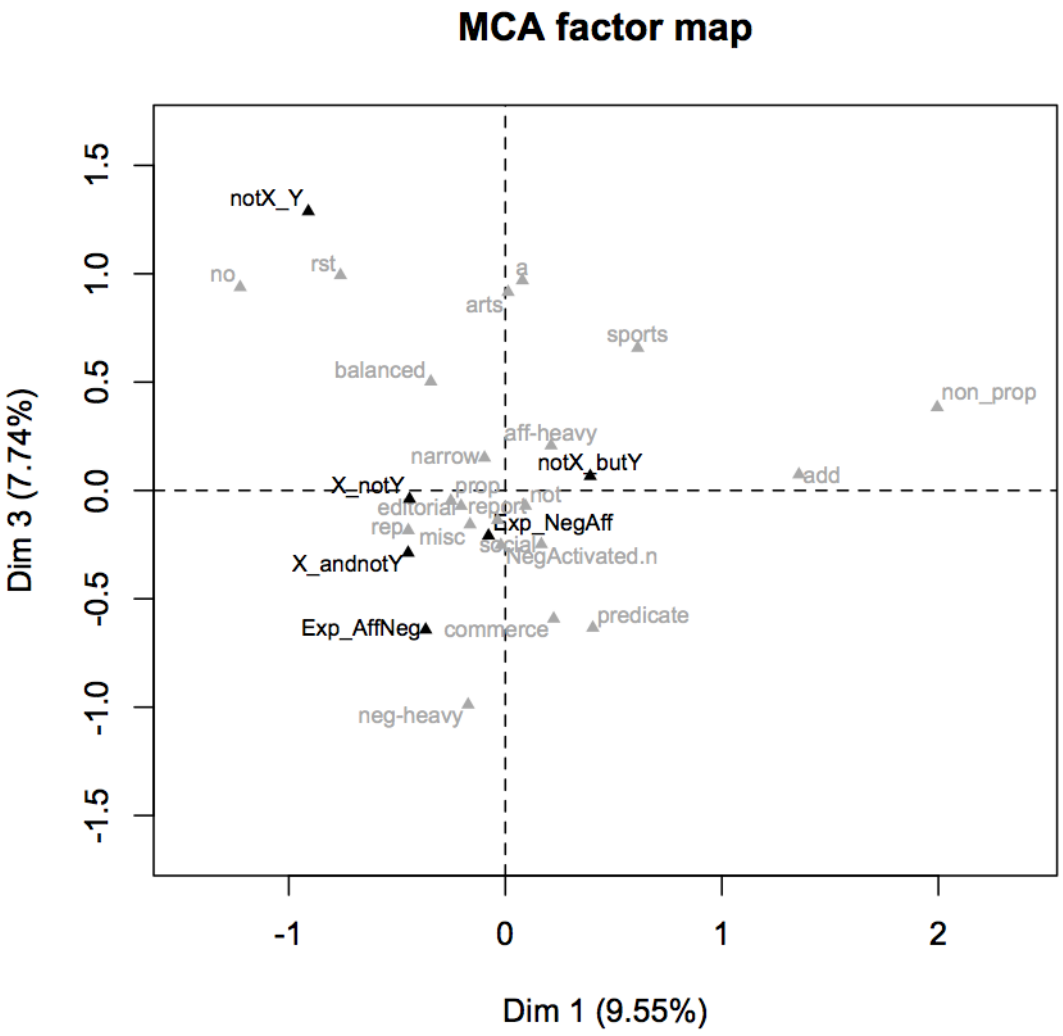
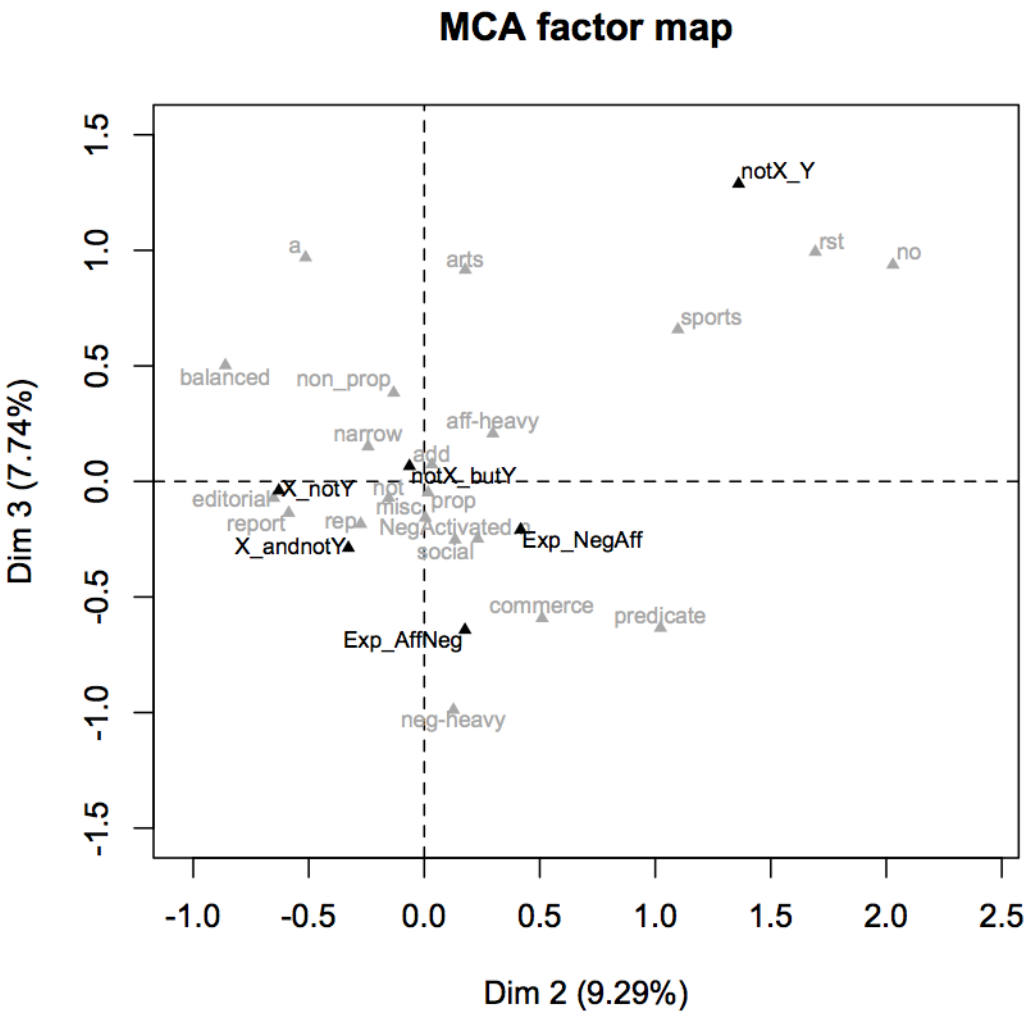


Figure 3. Biplot of dimensions 2 and 3



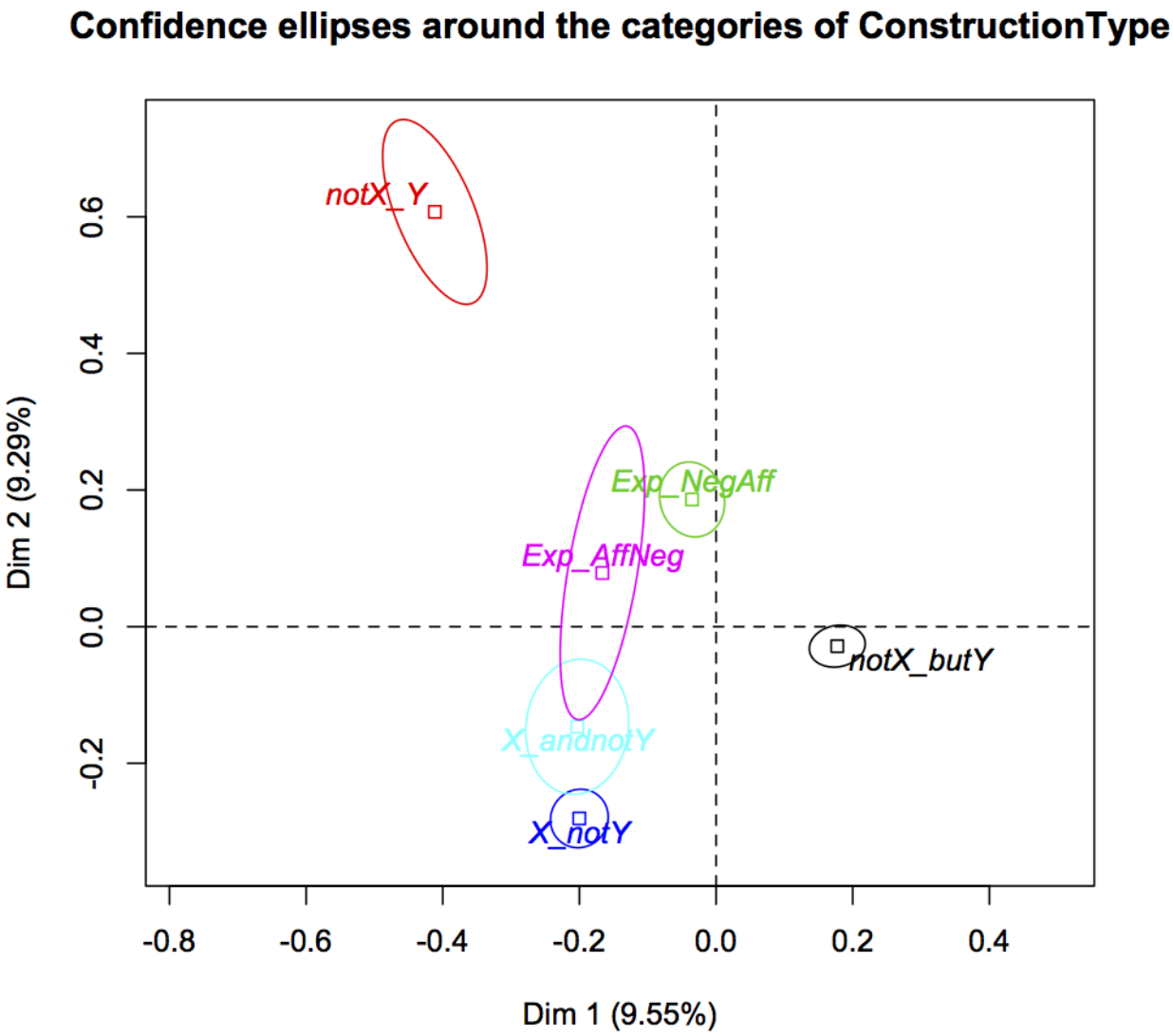
According to the variable contributions in Table 5, dimension 3 is predominantly based on WEIGHT, GENRE and NEGACTIVATED. The positive end is associated with the following features: activated negative foci, the genre of arts and affirmative foci that are equally long as or, less importantly for this dimension, longer than the contrasting negative foci. Correspondingly, non-activated negative foci and negative foci longer than their affirmative pairs cluster towards the negative end of the pole, as do certain written genres like commerce and science texts. Thus, while dimension 2 pitted unbalanced contrasts on one side and balanced contrasts on the other, dimension 3 pairs balanced contrasts with neg-heavy ones against aff-heavy contrasts. None of the schemas have a particularly strong affinity with either end of the pole though [*not X but Y*] and the Expanded [*not X, Y*] reach

significance for the positive side of the dimension, and so do [*X not Y*] and [*X and not Y*] for the negative side.

What can MCA tell us about constructional synonymy? The starting-point for this study is that usage reflects the functions of a construction. Since MCA directly models usage, the biplots also allow us to measure the degree of synonymy among the schemas. This can be done by drawing confidence ellipses around the centroids of the schemas. The centroids represent the prototypes of each schema. The closer the centroids, the more similar the schemas are. I begin by considering dimensions 1 and 2, in Figure 4. The figure shows that the three negative-second schemas are not clearly differentiated: there is a gradient along dimension 2 starting from the Expanded [*X not Y*], which overlaps with [*X and not Y*], which in turn borders [*X not Y*]. The other schemas are clearly differentiated, although the Expanded [*not X, Y*] is quite close to Expanded [*X not Y*].

@@ Insert Figure 4 here

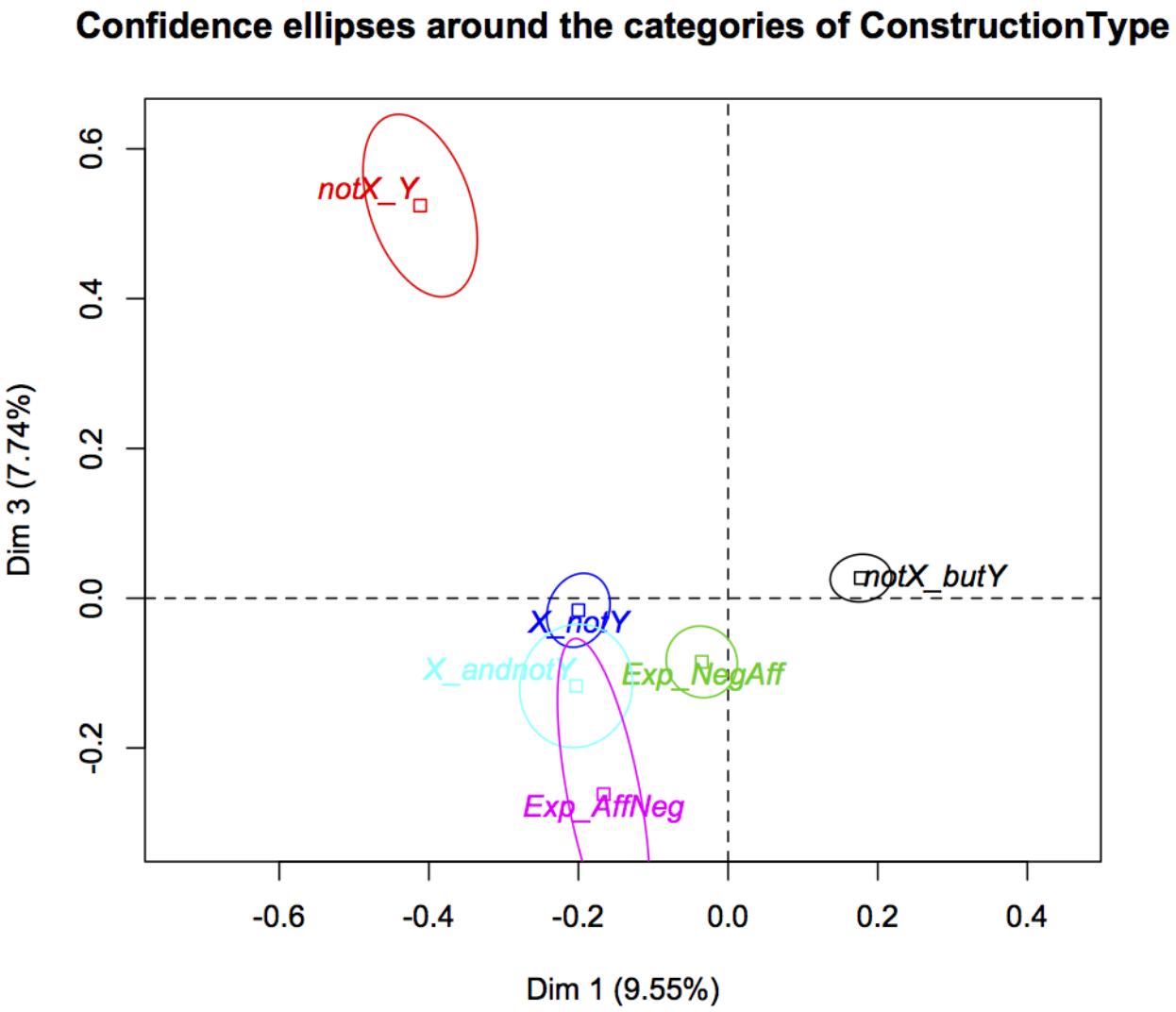
Figure 4. Confidence ellipses around the schemas in dimensions 1 and 2



When we turn to a similar analysis of dimensions 1 and 3 in Figure 5, we can see that the negative-second schemas still overlap. The analysis of dimensions 2 and 3, which is not shown, brings no change. Therefore, there is no statistically significant difference in the usage patterns of Expanded [*X not Y*] and [*X and not Y*] as well as non-expanded [*X not Y*] and [*X and not Y*], a point to which I shall return below.

@@ Insert Figure 5 here

Figure 5. Confidence ellipses around the schemas in dimensions 1 and 3



The last step of the analysis is model validation. Levshina (2015:383–385) suggests that the validity of a MCA should be confirmed using logistic regression. In logistic regression, the behaviour of a categorical dependent variable is modelled based on independent variables. In this case, I use CONSTRUCTIONTYPE as the dependent variable and the three dimensions of the MCA as independent variables. Since CONSTRUCTIONTYPE is a variable with six levels, I used multinomial logistic regression (see, e.g., Arppe 2008), performed with the `mlogit` package (Croissant 2013). Full results of the logistic regression analysis are given in Appendix 2, but to summarise, McFadden's R^2 is approximately 0.16 for the model. Values between 0.2 and 0.4 would indicate an

excellent fit. Thus, the MCA accounts for much of the variation in this dataset, but it does leave some variation unaccounted for. This is expected, since there is overlap even among the cores of the schemas.[19]

5. Discussion and conclusion

This article has presented an exploratory study of constructional variation in the domain of contrastive negation using multiple correspondence analysis. The analysis uncovered substantial probabilistic differences among the members of the construction family. The expanded schemas seem to be the most neutral members of the construction family, and the rest of the schemas deviate from them in various ways.

I shall now return to the research questions. The first research question was "what usage patterns are associated with which constructional schemas?" In other words, what are the schemas like? Table 6 presents a summary of their differentiating characteristics based on the univariate, bivariate and multivariate findings of this study.

@@ Insert Table 6 here

Table 6. The properties of the constructions

Constructional schema	**Properties**
[*not X but Y*]	• Typically expresses both replacive and additive meanings • Additive constructs often implicature-denying
[*not X, Y*]	• Typically expresses both replacive and restrictive meanings • Appears with both *not*-negation and *no*-negation
Expanded [*not X, Y*]	• Few features distinguishing it from the other schemas
[*X not Y*]	• Typically expresses replacive meanings • Almost never with *no*-negation or predicate-focus • Typical with contrastive foci that are equally long
[*X and not Y*]	• Similar to [*X not Y*] though less differentiated

[19] A reviewer points out that logistic regression could also be used on the "raw" data itself, i.e., the original variables rather than the MCA dimensions. There are relatively few exemplars of some of the constructional schemas discussed in this study, especially Expanded [*X not Y*]. Since the number of variables included in a logistic regression must be proportional to the lowest number of observations of a category, the number of variables would be much lower than in this study, or the number of observations higher. Thus, a logistic regression analysis of the constructional variation in the domain of contrastive negation is left for future work.

Expanded [*X not Y*]	• Few distinguishing features, not clearly separate from the other negative-second schemas, particularly [*X and not Y*]

The second and third research questions – "how close are the schemas to one another in terms of usage patterns?" and "are there true synonyms among the schemas?" – can be answered on the basis of the multivariate statistical analysis. The analysis suggests that the schemas form four clusters: (i) the expanded [*not X, Y*], (ii) the negative-second schemas, (iii) the [*not X but Y*] schema, and (iv) the [*not X, Y*] schema. From the point of view of constructional synonymy, this study has shown that even in conceptually abstract grammatical domains, forms tend to shy away from true synonymy. The results thus broadly support the Principle of No Synonymy as a general organising force in the constructicon while suggesting that it may not be exceptionless, especially with rarer and possibly under-entrenched schemas such as the Expanded [*X not Y*]. Lower-level abstractions may tell a different story, however: fixed expressions such as *children should be seen and not heard* may be entrenched in their exact forms and allow less variability across schemas that are otherwise close to one another. What happens at lower levels of the contrastive negation constructicon is largely outside the scope of this paper, but such entrenchment would be in line with Cappelle's (2006) and Perek's (2015) accounts of constructional alternations, in which the synonymy of two alternating formal patterns is itself part of speakers' mental grammars.

It is worth exploring the profiles of the various schemas in some more detail. The Expanded [*not X, Y*] schema has few distinguishing features in the multivariate analysis. All levels of all variables are reasonably common with it, and it has no particularly strong associations with any of the dimensions uncovered in MCA. It thus seems to be a neutral choice among the constructions studied here.

The negative-second schemas form a continuum along dimensions 2 and 3, and are poorly if at all differentiated by dimension 1. The [*X and not Y*] schema in particular is rather indistinct from the

other negative-second schemas. Also the Expanded [*X not Y*] schema has relatively little area of its own in the biplots. Together with its rather low number of occurrences, this suggests that it is the schema least likely to be entrenched in speakers' mental constructicons. Rather, it may emerge instantaneously and be figured out, possibly as a result of parallelism (on handling parallelism in a construction-based framework, see Murphy 2006; Jones et al. 2012). In Sinclair's (1987) terms, the Expanded schemas may be more open to the open-choice principle, i.e., their structure would be produced piece by piece rather than as a single chunk. This is supported by the fact that both of them are rather underdifferentiated in terms of their usage features.[20] Further psycholinguistic research is needed on this point, however.

The two sub-clausal negative-second schemas appear not to differ much from one another on any of the variables considered here. Possibly the most interesting feature of [*X not Y*] is its propensity for balanced contrasts. This creates rhetorically powerful, slogan-like constructs (e.g., *Make love, not war!*). The [*X and not Y*] schema seems to be largely absent from such usages.

Otherwise, the most interesting results were obtained for [*not X but Y*] and [*not X, Y*]. They seem to have specialised for different semantic types, which also affects the type of negation in them: often, [*not X but Y*] has additive semantics, which pairs up naturally (though not obligatorily) with implicature denial, whereas [*not X, Y*] is strongly associated with restrictive semantics, which tends to come with proposition denial. In fact, in many of the restrictive examples of [*not X, Y*], the focusing adverb takes a position reminiscent of a conjunction, as in (35). Without diachronic evidence, it is impossible to say whether this represents incipient constructionalisation of a potential [*not X just Y*] construction (Traugott & Trousdale 2013) or if this is a stable construction in English grammar. However, *but* itself has seen such a change from restrictive adverb to a conjunction; a remnant of *but* as an adverb is seen is in such examples as *She is but a little child* (see, e.g., Nevalainen 1991:124–127). A similar change has been suggested to have happened in French, for

[20] I thank an anonymous reviewer for pointing this out.

instance (Charolles & Lamiroy 2007), where the adverbs *seulement* 'only' and *simplement* 'simply' can sometimes be used instead of the adversative conjunction *mais* 'but', as in (37):

(37) Cette maison n'a pas de jardin, seulement/simplement/mais elle a une terrasse.
'This house does not have a garden, only/simply/but it has a terrace.' (Charolles & Lamiroy 2007:106, modified)

If the focusing adverb takes the position of a conjunction in the [*not X, Y*] schema, would it not be better analysed as the [*not X but Y*] schema in which the position of *but* is schematic to allow other conjunctions or conjunction-like expressions in that slot? This would mean that the focusing adverbs have constructionalised into conjunctions. However, the findings of this paper caution against such a view. The schemas have quite divergent behaviour. For instance, [*not X but Y*] freely allows predicate-focus, while [*not X, Y*] seems to be repelled by it as strongly as [*X not Y*] is. Therefore, it is implausible that restrictive [*not X, Y*] constructs would inherit the English coordination construction or that restrictive [*not X, Y*] constructs would be formed as a result of full analogisation with it.

The differences enumerated above are mostly tendencies rather than absolutes. This supports the view that constructional synonymy is gradient. This study has utilised a method, multiple correspondence analysis, that readily captures this gradience while also allowing us to see the conceptual areas that show entrenched usage patterns. In the domain of contrastive negation, this is particularly the case of implicature-denying additives and the [*not X but Y*] schema on one hand, and restrictives and the [*not X, Y*] schema on the other hand. Further work is needed to ascertain whether and to what extent the [*not only X but also Y*] and [*not X, just Y*] constructions are cognitively linked to the constructional schemas posited here.

Finally, it is useful to consider the results from a typological perspective. English has no special markers for contrastive negation. Thus, it differs from languages which make a distinction between adversative and corrective coordination, both of which are expressed by English *but*. In languages

that do make such a distinction, corrective coordinators are used in constructions equivalent to [*not X but Y*] in lieu of *but*. Thus, German makes a distinction between adversative *aber* and corrective *sondern*, and only the latter is used in [*nicht X sondern Y*] 'not X but Y' (Anscombre & Ducrot 1977; Horn 1985; 1989). Another relevant parameter is whether a language has negators specialised for contrastive negation, such as *non pas* in French, which can appear in many kinds of constructions (e.g., [*non pas X mais Y*] 'not X but Y' or [*X non pas Y*] 'X not Y'). Future cross-linguistic studies will hopefully shed light on whether languages like German and French, which have carved up an overtly marked construction of contrastive negation, partition the conceptual space differently from languages like English, which have not done so.

Acknowledgements

Many people have helped me in the process of writing this article. I particularly wish to thank Matti Miestamo, Minna Palander-Collin, Tanja Säily and Jouni Rostila for their comments, which have helped me to improve this article considerably. Kimmo Vehkalahti deserves my gratitude for originally suggesting MCA as a method to me. I also wish to thank the anonymous reviewers and the editors for their constructive feedback on an earlier version. I alone am responsible for all shortcomings of the study.

References

Anscombre, J.-C., & Ducrot, O. (1977). Deux *mais* en français? *Lingua* 43(1), 23–40.

Arppe, A. 2008. Univariate, bivariate and multivariate methods in corpus-based lexicography: a study of synonymy. Doctoral dissertation, University of Helsinki. Helsinki: Helsinki University Print. https://helda.helsinki.fi/handle/10138/19274.

Baayen, R. H. (2008). *Analyzing Linguistic Data: A Practical Introduction to Statistics using R*. Cambridge: Cambridge University Press.

Behaghel, O. (1909/1910). Beziehungen zwischen Umfang und Reihenfolge von Satzgliedern. *Indogermanische Forschungen* 25, 110–142.

Berglund, Y., Hoffmann, S., Lee, D. & Smith, N. (2002). BNCweb manual. http://rdues.bcu.ac.uk/bncweb/manual/bncwebman-home.htm.

Blanco, E., & Moldovan, D. (2013). Retrieving implicit positive meaning from negated statements. *Natural Language Engineering* 20(4), 501–535. doi:10.1017/S1351324913000041.

BNC = *British National Corpus* (1991–1994). BNC Consortium. http://www.natcorp.ox.ac.uk.

Bolinger, D. (1977). *Meaning and Form*. London & New York: Longman.

Bresnan, J., Cueni A., Nikitina, T., & Baayen, R. H. (2007). Predicting the dative alternation. In G. Bouma, I. Krämer & J. Zwarts (Eds.), *Cognitive Foundations of Interpretation* (pp. 69–94). Amsterdam: Royal Netherlands Academy of Science.

Cappelle, B. (2006). Particle placement and the case for "allostructions." *Constructions* 8 (Special volume 1).

Carston, R. (1996). Metalinguistic negation and echoic use. *Journal of Pragmatics* 25(3), 309–330.

Carston, R. (1998). Negation, "presupposition" and the semantics/pragmatics distinction. *Journal of Linguistics* 34(2), 309–350.

Charolles, M., & Lamiroy, B. (2007). Du lexique à la grammaire : *seulement*, *simplement*, *uniquement*. *Cahiers de lexicologie* 90, 93–116.

Chatar-Moumni, N. (2008). Quelques aspects du fonctionnement de la négation en arabe marocain. *La Linguistique* 44(2), 81–98.

Croissant, Y. (2013). mlogit: multinomial logit model. https://cran.r-project.org/package=mlogit.

Dik, S., Hoffmann, M. E., de Jong, J. R., Sie Ing Djiang, Stroomer, H., & de Vries, L. (1981). On the typology of focus phenomena. In T. Hoekstra, H. van der Hulst & M. Moortgat (Eds.), *Perspectives on Functional Grammar* (pp. 41–74). Dordrecht: Foris Publications.

Fillmore, C. J., Kay, P., & O'Connor, M. C. (1988). Regularity and idiomaticity in grammatical constructions: The case of *let alone*. *Language* 64(3), 501–538.

Gates Jr., D. L., & Seright, O. D. (1967). Negative-contrastive constructions in standard modern English. *American Speech* 42(2), 136–141. doi:10.2307/453965.

Geurts, B. (1998). The mechanisms of denial. *Language* 74(2), 274–307.

Givón, T. (1978). Negation in language: Pragmatics, function, ontology. In P. Cole (Ed.), *Syntax and Semantics 9: Pragmatics* (pp. 69–112). New York: Academic Press.

Glynn, D. (2014). Correspondence analysis: Exploring data and identifying patterns. In D. Glynn & J. A. Robinson (Eds.), *Corpus Methods for Semantics: Quantitative Studies in Polysemy and Synonymy* (pp. 443–485). Amsterdam/Philadelphia: John Benjamins.

Goldberg, A. E. (1995). *Constructions: A Construction Grammar Approach to Argument Structure*. Chicago: University of Chicago Press.

Goldberg, A. E. (2006). *Constructions at Work: The Nature of Generalization in Language*. Oxford: Oxford University Press.

Granvik, A., & Taimitarha, S. (2014). Topic-marking prepositions in Swedish: A corpus-based analysis of adpositional synonymy. *Nordic Journal of Linguistics* 37(2), 257–296.

Greenacre, M. (2017). *Correspondence Analysis in Practice*. Third edition. Boca Raton, FL: Taylor & Francis.

Gries, S. Th. (2003). *Multifactorial Analysis in Corpus Linguistics: A Study of Particle Placement*. New York & London: Continuum.

Hawkins, J. A. (2003). Efficiency and complexity in grammars: Three general principles. In J. Moore & M. Polinsky (Eds.), *The Nature of Explanation in Linguistic Theory* (pp. 121–152). Stanford: CSLI Publications.

Hilpert, M. (2015). From *hand-carved* to *computer-based*: Noun-participle compounding and the upward strengthening hypothesis. *Cognitive Linguistics* 26(1), 113–147. doi:10.1515/cog-2014-0001.

Hoffmann, S., Evert, S., Smith, N., Lee, D., & Berglund Prytz, Y. (2008). *Corpus Linguistics with BNCweb – a Practical Guide*. Frankfurt: Peter Lang.

Horn, L. R. (1985). Metalinguistic Negation and Pragmatic Ambiguity. *Language* 61(1), 121–174. doi:10.2307/413423.

Horn, L. R. (1989). *A natural history of negation*. Chicago: University of Chicago Press.

Huddleston, R., & Pullum, G. K. (2002). *The Cambridge Grammar of the English Language*. Cambridge: Cambridge University Press.

Jones, S., Murphy, M. L., Paradis, C., & Willners, C. (2012). *Antonyms in English: Construals, Constructions and Canonicity*. Cambridge: Cambridge University Press.

Kroeger, P. (2014). External negation in Malay/Indonesian. *Language* 90(1), 137–184.

Lakoff, G. (1987). *Women, Fire, and Dangerous Things: What Categories Reveal about the Mind*. Chicago: University of Chicago Press.

Lambrecht, K. (1994). *Information structure and sentence form: Topic, focus, and the mental representations of discourse referents*. Cambridge: Cambridge University Press.

Langacker, R. W. (1987). *Foundations of Cognitive Grammar, Volume 1: Theoretical Prerequisites*. Stanford: Stanford University Press.

Lê, S., Josse, J., & Husson, F. (2008). FactoMineR: An R Package for Multivariate Analysis. *Journal of Statistical Software* 25(1). doi:10.18637/jss.v025.i01. https://www.jstatsoft.org/article/view/v025i01.

Lee, D. Y. W. (2001). Genres, registers, text types, domains and styles: Clarifying the concepts and navigating a path through the BNC jungle. *Language Learning & Technology* 5(3), 37–72. http://llt.msu.edu/vol5num3/pdf/lee.pdf.

Le Roux, B., & Rouanet, H. (2010). *Multiple Correspondence Analysis*. Los Angeles, CA: SAGE. doi:10.4135/9781412993906.

Levshina, N. (2015). *How to do Linguistics with R: Data exploration and statistical analysis*. Amsterdam/Philadelphia: John Benjamins. doi:10.1075/z.195.

McCawley, J. D. (1991). Contrastive Negation and Metalinguistic Negation. *CLS* 27(2), 189–206.

Mughazy, M. (2003). Metalinguistic negation and truth functions: the case of Egyptian Arabic. *Journal of Pragmatics* 35(8), 1143–1160.

Murphy, M. L. (2006). Antonyms as lexical constructions: or, why *paradigmatic construction* is not an oxymoron. *Constructions* 8 (Special volume 1).

Nenadic, O., & Greenacre, M. (2007). Correspondence Analysis in R, with Two- and Three-dimensional Graphics: The ca Package. *Journal of Statistical Software* 20(3). doi:10.18637/jss.v020.i03. https://www.jstatsoft.org/article/view/v020i03.

Nevalainen, T. (1991). *BUT, ONLY, JUST: Focusing adverbial change in Modern English 1500–1900.* Helsinki: Société Néophilologique.

Paolillo, J. C. (2002). *Analyzing Linguistic Variation: Statistical Models and Methods*. Stanford: CSLI Publications.

Perek, F. (2015). *Argument Structure in Usage-Based Construction Grammar: Experimental and Corpus-Based Perspectives*. Amsterdam/Philadelphia: John Benjamins.

Prince, E. F. (1981). Toward a Taxonomy of Given-New Information. In P. Cole (Ed.), *Radical Pragmatics* (pp. 223–255). New York: Academic Press.

Quirk, R., Greenbaum, S., Leech, G., & Svartvik, J. (1985). *A Comprehensive Grammar of the English Language*. Harlow: Longman.

R Core Team. (2016). R: A language and environment for statistical computing. Vienna: R Foundation for Statistical Computing.

Sheskin, D. J. (2011). *Handbook of Parametric and Nonparametric Statistical Procedures*. Boca Raton, FL: Chapman and Hall/CRC Press.

Silvennoinen, O. O. (2017). Not only apples but also oranges: Contrastive negation and register. In T. Hiltunen, J. McVeigh & T. Säily (Eds.), *Big and Rich Data in English Corpus Linguistics: Methods and Explorations*. Helsinki: VARIENG. http://www.helsinki.fi/varieng/series/volumes/19/silvennoinen/.

Sinclair, J. McH. (1987). Collocation: A progress report. In R. Steele & T. Threadgold (Eds.), *Language Topics: Essays in Honour of Michael Halliday* (pp. 319–331). Amsterdam/Philadelphia: John Benjamins.

Szmrecsanyi, B. (2004). On operationalizing syntactic complexity. In G. Purnelle, C. Fairon & A. Dister (Eds.), *Le poids des mots. Proceedings of the 7th International Conference on Textual Data Statistical Analysis. Louvain-la-Neuve, March 10-12, 2004, Vol. 2* (pp. 1032–1039). Louvain-la-Neuve: Presses universitaires de Louvain.

Szmrecsanyi, B. (2006). *Morphosyntactic Persistence in Spoken English: A Corpus Study at the Intersection of Variationist Sociolinguistics, Psycholinguistics, and Discourse Analysis*. Berlin/New York: Mouton de Gruyter.

Toosarvandani, M. D. (2010). Association with Foci. Doctoral dissertation, University of California, Berkeley.

Tottie, G. (1991). *Negation in English Speech and Writing: A Study in Variation*. San Diego: Academic Press.

Traugott, E. C., & Trousdale, G. (2013). *Constructionalization and Constructional Changes*. Oxford: Oxford University Press.

Van Valin Jr., R. D. (2005). *Exploring the Syntax-Semantics Interface*. Cambridge: Cambridge University Press.

Wasow, T. (1997). Remarks on grammatical weight. *Language Variation and Change* 9(1), 81–105. doi:10.1017/S0954394500001800.

Weller, S. C., & Romney, K. (1990). *Metric Scaling: Correspondence Analysis*. Newbury Park, CA: SAGE.

Yeh, L. (1995). Focus, metalinguistic negation and contrastive negation. *Journal of Chinese Linguistics* 23(2), 42–75.

Appendices

Appendix 1. Associations among variables

The upper half of the table shows the significance levels of the p-values while the lower half presents the chi square value and the degrees of freedom. The significance levels: *** = very highly statistically significant ($p < 0.001$), ** = highly statistically significant ($p < 0.01$), * = statistically significant ($p < 0.05$), — = not statistically significant ($p > 0.05$).

	ConstructionType	Neg	SemType	Weight	NegTarget	NegActivated	FocusStr	Genre
ConstructionType		***	***	***	***	—	***	(*)[21]
Neg	152.41 df = 5		***	—	—	—	*	*
SemType	799.21 df = 10	124.89 df = 2		***	***	—	*	**
Weight	283.05 df = 10	5.6603 df = 2	38.55 df = 6		—	*	***	*
NegTarget	26.445 df = 5	2.9624 df = 1	221.43 df = 2	3.3983 df = 2		—	—	**
NegActivated	10.352 df = 5	0.57003 df = 1	0.5556 df = 2	9.245 df = 2	0.58553 df = 1		**	*
FocusStr	331.6 df = 5	8.6625 df = 1	10.252 df = 2	44.577 df = 2	1.3119 df = 1	14.701 df = 1		*
Genre	(61.018) (df = 35)	15.521 df = 7	36.592 df = 14	25.609 df = 14	4.8652 df = 7	15.195 df = 7	14.818 df = 7	

Appendix 2. Model validation using logistic regression

```
Call:
mlogit(formula = ConstructionType ~ 1 | Dim1 + Dim2 + Dim3, data =
cn3,
    reflevel = 1, method = "nr", print.level = 0)
```

[21] There exist various heuristics for deciding which significance test to use. The strictest view is that one should never use the chi square test when the dataset includes cells with expected values below 5, a criterion that my dataset fails. According to a laxer criterion, when the total number of observations is over 40, the chi square test is acceptable if no expected value goes below 1 (Sheskin 2011, cited in Levshina, 2015, p. 214). This criterion is met, except for the association between ConstructionType and Genre; here, the expected frequency of Expanded [*X not Y*] in the science subcorpus is approximately 0.97. This is why the respective information in the table is in parentheses. The alternative to chi square, Fisher's Exact, is computationally unfeasible for my dataset. Therefore, the chi square test was chosen.

```
Frequencies of alternatives:
Exp_AffNeg Exp_NegAff  notX_butY     notX_Y  X_andnotY     X_notY
  0.015838   0.250683   0.444020   0.043692   0.054069   0.191699

nr method
6 iterations, 0h:0m:1s
g'(-H)^-1g = 2.18E-07
gradient close to zero

Coefficients :
                         Estimate Std. Error t-value  Pr(>|t|)
Exp_NegAff:(intercept)   3.041531   0.264134 11.5151 < 2.2e-16 ***
notX_butY:(intercept)    3.654611   0.261755 13.9619 < 2.2e-16 ***
notX_Y:(intercept)       0.192237   0.339828  0.5657 0.5716044
X_andnotY:(intercept)    1.266822   0.295293  4.2901 1.786e-05 ***
X_notY:(intercept)       2.388619   0.271900  8.7849 < 2.2e-16 ***
Exp_NegAff:Dim1          1.328778   0.664790  1.9988 0.0456307 *
notX_butY:Dim1           2.369993   0.662891  3.5752 0.0003499 ***
notX_Y:Dim1             -0.489846   0.798055 -0.6138 0.5393474
X_andnotY:Dim1          -0.303421   0.771456 -0.3933 0.6940907
X_notY:Dim1              0.012763   0.693202  0.0184 0.9853105
Exp_NegAff:Dim2          0.617485   0.463131  1.3333 0.1824388
notX_butY:Dim2          -0.637040   0.465602 -1.3682 0.1712473
notX_Y:Dim2              0.861527   0.520644  1.6547 0.0979786 .
X_andnotY:Dim2          -1.415786   0.538824 -2.6275 0.0086002 **
X_notY:Dim2             -2.531275   0.498332 -5.0795 3.784e-07 ***
Exp_NegAff:Dim3          1.214278   0.492587  2.4651 0.0136975 *
notX_butY:Dim3           1.845244   0.492541  3.7464 0.0001794 ***
notX_Y:Dim3              4.147457   0.609270  6.8073 9.948e-12 ***
X_andnotY:Dim3           0.785753   0.555759  1.4138 0.1574094
X_notY:Dim3              1.283059   0.510706  2.5123 0.0119939 *
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Log-Likelihood: -2140.6
McFadden R^2:  0.15538
Likelihood ratio test : chisq = 787.58 (p.value = < 2.22e-16)
```