



HAL
open science

Visualizing distances in a set of near synonyms: rather, quite, fairly, and pretty.

Guillaume Desagulier

► To cite this version:

Guillaume Desagulier. Visualizing distances in a set of near synonyms: rather, quite, fairly, and pretty.. Corpus Methods for Semantics: Quantitative studies in polysemy and synonymy, Human Cognitive Processing 43, John Benjamins, pp.145-178, 2014, 9789027223975. 10.1075/hcp.43.06des . halshs-00731035v2

HAL Id: halshs-00731035

<https://shs.hal.science/halshs-00731035v2>

Submitted on 7 Dec 2017

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

prepublication copy

Visualizing distances in a set of near-synonyms

Rather, quite, fairly, and pretty

Guillaume Desagulier

Université Paris 8, Université Paris Ouest /
Nanterre La Défense, UMR 7114 MoDyCo

I aim to uncover revealing aspects of the conceptual structure of four English moderators (*rather, quite, fairly, and pretty*) and shed new light on previous studies on degree modifiers. I develop an original methodology for handling and visualizing measures of significant attraction between lexical items. This methodology combines univariate and multivariate statistics. Collexeme analysis is used as input for hierarchical agglomerative cluster analysis, and multiple distinctive collexeme analysis is used as input for correspondence analysis. Visualizing collexemes with exploratory tools does more than depict proximities and distances between individuals and variables. It is also an accurate means to (a) unveil fine semantic differences in a set of near-synonymous constructions, (b) determine entrenchment continua, and (c) represent a significant part of the complex inventory of intensifying constructions.

Keywords: Cognitive Construction Grammar, collostructional analysis, correspondence analysis, degree modifiers, hierarchical cluster analysis, near-synonymy

1. Introduction

It is a well-known fact that natural languages avoid true synonymy: “languages abhor absolute synonyms just as nature abhors a vacuum” (Cruse 1986:270). This is why absolute synonyms are rare whereas near-synonyms are extremely frequent. In Cognitive-Grammar terms (Langacker 1987; 1991), synonymous expressions have identical conceptual content and impose the same construal upon that conceptual content, while near-synonyms share the same conceptual content but differ in terms of

construal.¹ If we could measure conceptual content similarities and construal differences, and then represent them graphically, we could unveil fine semantic differences in sets of near-synonyms.

Bearing in mind that “a word is known by the company it keeps” (Firth 1957), I explore the collocation preferences of four English degree modifiers: *rather*, *quite*, *pretty*, and *fairly*. Because these adverbs are near-synonyms, we may expect them to share identical conceptual content but differ in how this conceptual content is construed. Although these adverbs can grade other adverbs (*pretty badly*, *fairly easily*) or noun phrases (*quite a shock*, *rather a surprise*), I restrict my investigation to the prototypical contexts where they are used as degree modifiers of adjectives, as examples (1) to (4) illustrate:

- (1) *I heard a rather odd conversation.*
(Corpus of Contemporary American English)
- (2) *These two teams are quite similar in some ways.* (*ibid.*)
- (3) *I think it's fairly easy for anyone to get anything they want (...).* (*ibid.*)
- (4) *He seemed in pretty good form.* (*ibid.*)

In the above sentences, *rather*, *quite*, *pretty*, and *fairly* index the properties of the adjectives they modify as ‘not fully X’. They function as ‘word modifiers’, not as ‘phrasal modifiers’ (Stoffel 1901). They are constituents of adjective phrases and they scale inherent qualities or properties of the heads.²

This paper addresses three challenges. The first challenge is to determine the conceptual content that *rather*, *quite*, *pretty*, and *fairly* presumably share. The second challenge is to spot the distinctive construals that these adverbs impose on the conceptual content of the adjectives they modify. The third challenge is to identify how adjectives modify the conceptual content of moderators. The working hypothesis is that overlap in collocation preferences will reveal that *rather*, *quite*, *fairly*, and *pretty* have similar conceptual content, whereas subtle differences will indicate that these four degree modifiers impose different modes of construal.

In the pages that follow, moderators are contrasted in their collocation preferences in the 415M-word Corpus of Contemporary American English (Davies 1990-present). The proposed method requires extracting all tokens of the <moderator +

1. For example, the distributive quantifiers *each* and *every* both particularize a referent by referring to its discreteness in a group. But they impose different ways of construing this conceptual content: *each* profiles a discrete entity within a group, whereas *every* profiles a group of discrete entities.

2. Moderators, and degree modifiers in general, should not be mixed up with scalar focus modifiers (*even*, *rarely*, *barely*). While the former are inherently scalar, focus modifiers are not: they merely evoke a scale (Traugott 2008).

adjective> construction and implementing two techniques from a family of methods known as collocation analysis. First, a collexeme analysis (Stefanowitsch & Gries 2003) determines which adjectives are most strongly attracted to each moderator. Overlap in collocation preferences will be an indication that moderators have similar conceptual content. Then, a multiple distinctive collexeme analysis (Gries & Stefanowitsch 2004) determines those adjectives that are distinctively associated with each moderator. If, despite overlap, the moderators attract distinct adjective classes, then: (a) moderators form a functionally coherent paradigm; (b) this paradigm has a complex internal structure.

When one investigates semantic similarities and differences between related lexemes or constructions, one can choose among many association measures, such as mutual information (*pointwise MI*, MI^2 , MI^3), the *t*-score, the Fisher-Yates exact test, the binomial test, the χ^2 test, etc. Not all of them are equally satisfactory though (Evert 2004; Wiechmann 2008). For example, pointwise mutual information (Church & Hanks 1990) tends to overestimate rare words and is sensitive to data sparseness (Manning and Schütze 1999: 180–181; Kilgarriff 2001). The χ^2 test presupposes that the linguistic phenomenon under investigation is distributed randomly across a corpus, but as Kilgarriff (2005) puts it, “language is never, ever, ever, random”.

Collocation analysis is a good place to start because it is a “safe” option that is based on association measures that (a) do not overestimate effect for low-frequency pairs and (b) do not violate distributional assumptions. However, one of the collocation analysis methods, namely multiple distinctive collexeme analysis, should be used cautiously. Multiple distinctive collexeme analysis is best used to investigate collocation preferences within a predefined closed set of functionally similar constructions (i.e. alternations). Therefore, one needs to make sure that the paradigm under investigation is structured coherently before running such an analysis. As far as moderators are concerned, this is no easy task because they can scale upwards or downwards depending on the adjectives they modify (Paradis 1997: 87). Additionally, past research has shown that context dependency between adverbs and adjectives is not always decisive (Allerton 1987; Paradis 1997; Athanasiadou 2007). It is often impossible to decide whether *quite* is a maximizer or a moderator. A second difficulty is to generalize the output of collocation analyses. Collexeme and (multiple) distinctive collexeme analyses output large tables containing as many rows as there are collexemes for each variable. The linguist uses these tables to find out whether collexemes fall into distinct semantic classes, but such classes are generally scattered across many rows, and are invisible to the naked eye. Finally, in order to fully determine the semantic profile of the <moderator + adjective> construction, one needs to capture not only the semantic relations between each moderator and its distinctive adjectives (which collocation analysis does), but also the semantic relations between moderators *and* between adjectives. A multifactorial analysis can handle all three types of semantic relations.

To solve each of these issues, the output of collexeme and multiple distinctive analyses serves as input for two multifactorial methods: hierarchical cluster analysis (Divjak & Fieller, this volume 405–442) and correspondence analysis (Benzécri 1973; Benzécri 1984; Greenacre 2007). Both are unsupervised, exploratory clustering techniques. They help find structure in multivariate data thanks to observation groupings. In both cases, the linguist makes no assumption as to what groupings should be there.

Hierarchical cluster analysis uses distance matrices to cluster data in a tree-like format – more specifically a dendrogram. This method is used to determine the preferred collexemes of 23 degree modifiers in English and to cluster the results to see if the classes that one obtains are consistent with the traditional functional classes that degree modifiers fall into: maximizers, boosters, moderators, approximators and diminishers (Allerton 1987; Paradis 1997; Quirk *et al.* 1985). If moderators form a coherent functional category, it is relevant to perform a multiple distinctive collexeme analysis. If this category proves coherent, it will come precisely from the kind of collocates that moderators attract.

To account for the semantic correspondences between moderators and adjectives, I focus on the distinctive lexical preferences of the four moderators *rather*, *quite*, *fairly*, and *pretty*. More specifically, I submit the output of multiple distinctive collexeme analysis to correspondence analysis, a multifactorial approach that provides a low-dimensional map of the data by calculating matrices between the rows and the columns of a contingency table. The larger the distance is between two rows or columns, the further apart the row or column coordinates will be on the map. Correspondence analysis has two advantages. First, it captures semantic relations between (a) moderators, (b) adjectives, and (c) moderators and adjectives. Second, it offers an elegant and efficient way of visualizing distances between lexical variables and collexemes, even when these are computed from large tables.

The paper is organized as follows. Section 2 reviews previous qualitative and quantitative research on degree modifiers of adjectives. Section 3 presents the corpus and the methods used. Section 4 summarizes the results. In Section 5, I discuss the results and explain why an original combination of univariate and multivariate statistical methods can only improve our understanding of near-synonymy.

2. Previous research

2.1 Typologies of degree modifiers

Degree modifiers – also known as intensifiers – are a subclass of degree words (Bolinger 1972). They give specifications of degree concerning the adjectives they modify. Adverbs such as *very*, *extremely*, *absolutely* scale adjectival properties “upwards”, whereas other adverbs, such as *slightly*, *a little*, *somewhat* scale adjectival

properties “downwards”. *Rather*, *quite*, *fairly*, and *pretty* set the qualities that gradable adjectives denote to a moderate level. Along with *moderately* and *relatively*, these degree modifiers are known as ‘moderators’ (Paradis 1997).

Like most degree modifiers, *rather*, *quite*, *fairly*, and *pretty* are typologically unstable because they do not always neatly fit in the functional categories that linguists have assigned them. For example, *quite* is likely to be interpreted as a maximizer when it modifies an extreme/absolute adjective (*this novel is quite excellent*) or a telic/limit/liminal adjective (*quite sufficient*), but it is likely to be a moderator when it modifies a scalar adjective (*quite big*) (Paradis 1997:87). Past research has shown that context dependency between adverbs and adjectives is not always decisive. It is often impossible to decide whether *quite* is a maximizer or a moderator. For example, *quite* is ambiguous when it modifies the adjective *different* (Allerton 1987:25). Recent claims that *quite* has undergone grammaticalization have led linguists to think that this adverb is ambiguous with absolute and scalar adjectives alike (*the village is quite beautiful*, *the play is quite good*). Similarly, *rather*, *pretty* and *fairly* can scale upwards or downwards due to increasing subjectification (Nevalainen & Rissanen 2002; Athanasiadou 2007).

Given that context dependency is the only way to determine whether a given degree modifier such as *quite* is a maximizer or a moderator, conducting a quantitative study of collocational preferences between adverbs and adjectives across corpora is a methodologically sound approach. It is also a popular approach, as evidenced by the number of corpus-based studies of degree modifiers from the early 1990s onwards (Altenberg 1991; Lorenz 2002; Kennedy 2003; Simon-Vandenberg 2008). Most of these quantitative approaches are intuitively attractive, but they are problematic for two reasons. First, some of them do not rely on corpora that are sufficient in size and thus fail to provide a detailed picture of the collocational preferences of degree modifiers. By way of illustration, the London Lund Corpus, which Altenberg (1991) and Paradis (1997) exploit, contains 0.5M-words. Some linguists could object that it is presumably too small to reveal relatively rare, but not necessarily statistically insignificant, patterns of attraction between intensifiers and adjectives. Second, linguists tend to use measures that are inadequate to reveal two-way interactions between collocants. Raw frequencies and coarse-grained relative frequencies such as percentages and counts per *n*-thousand words (Altenberg 1991; Paradis 1997) underestimate significant collocations because they do not filter away lexemes that are highly frequent regardless of the specific contexts where they occur. Likewise, the choice of pointwise mutual information (Kennedy 2003) is not particularly helpful for the reasons outlined in Section 1.³

3. This can be confirmed easily on COCA, whose search interface allows end users to compute MI scores: (corpus.byu.edu/mutualinformation.asp).

In the wake of Stoffel (1901), typologies of degree modifiers have proliferated (Quirk *et al.* 1985; Allerton 1987; Paradis 1997), a sign that these adverbs do not all lend themselves easily to a classificatory exercise. One recurring issue is the lack of a fully satisfactory organizing principle. Quirk *et al.* (1985) posit a coarse-grained distinction between ‘amplifiers’, which “scale upward from an assumed norm”, and ‘downtoners’, which “scale downwards”.⁴ Amplifiers further subdivide into ‘maximizers’ and ‘boosters’, and downtoners into ‘approximators’, ‘compromisers’, ‘diminishers’, and ‘minimizers’. However, Quirk *et al.*’s classification is problematic in at least two respects. Firstly, its internal structure is unjustified and inconsistent. In this regard, Allerton (1987: 18–19) and Paradis (1997: 24) observe that the distribution of degree modifiers across the aforementioned categories is inaccurate. Secondly, Quirk *et al.* disregard the fact that the semantic influence between an intensifier and the head it modifies is bidirectional.

Allerton observes that some approximators do not occur with all adjectives (e.g. *virtually unique* vs. *?virtually large*). He proposes a four-entry classification of degree modifiers of adjectives (scalar modifiers, telic modifiers, absolute modifiers, differential modifiers). This classification takes into account the collocational restrictions of the <degree modifier + adjective> sequence (1987: 19–21). However, Allerton gives no corpus evidence to support his claims and, although convincing, his model is not empirically grounded.

Paradis’s taxonomy is finer-grained. Like Quirk *et al.*, Paradis (1997: 27–28) proposes a general bipartition between modifiers that scale upwards (‘reinforcers’) and modifiers that scale downwards (‘attenuators’). She postulates the existence of a cline between these two poles. She unifies each set on the cline by means of the principle of cognitive synonymy (Cruse 1986). She further subdivides ‘reinforcers’ and ‘attenuators’ into ‘totality modifiers’ and ‘scalar modifiers’. Unlike Quirk *et al.*, but like Allerton, Paradis is well aware that some degree modifiers are polysemous, semantically fuzzy, and therefore paradigmatically unstable. Unlike Allerton’s, Paradis’s typology is empirically founded. She devotes a chapter of her monograph to the distribution of degree modifiers of adjectives across the London-Lund Corpus of spoken English (1997: Chapter 2). Her goal is both to examine the distribution of degree modifiers of adjectives across modes (spoken vs. written) and to inspect the detail of collocational preferences in the hope of establishing a subtle classification. She concludes that intonation is the key to understanding the meaning and the grading force of degree modifiers (1997: Chapter 4).

Previous corpus-based approaches to degree modifiers make extensive use of raw counts, or relative frequencies that do not filter away overrepresented adjectives. Other

4. Strictly speaking, Quirk *et al.*’s two-branch model is restricted to degree modifiers used as subjuncts, but as Paradis (1997: 23–24) observes, this model applies equally well to degree modifiers of adjectives.

approaches may also suffer from the kind of corpus that they use, the lack of statistical methods, or the selection of inadequate statistics. If one uses a corpus that is relatively small, statistical methods that are invalid for low counts, such as (pointwise) mutual information or χ^2 , are inappropriate, and one ends up being trapped in a vicious circle. This is a problem if one wants to investigate collocations, some of which might be relatively rare but by no means insignificant, both statistically and semantically.

2.2 Corpus-based Cognitive Linguistics

Cognitive Linguistics is a usage-based approach to language that makes no principled distinction between language use and language structure. A linguistic unit is entrenched and stored in grammar when a usage pattern generalizes across recurring instances of language use. The more frequently speakers encounter a linguistic unit, the more the linguistic unit is entrenched, i.e. established as a cognitive routine. In Cognitive Linguistics, entrenched patterns of usage provide privileged access to speakers' knowledge of their language.

The meaning of a linguistic unit “involves both conceptual content and the construal of that content” (Langacker 2008:44). Conceptual content is referred to as a domain: a consistent knowledge representation that serves as a basis for the construal of other conceptual units. For instance, adjectives such as *damp*, *dampish*, *moist* or *dry* relate to conceptual units to be construed with respect to the domain of WETNESS. The content that the domain designates can further be construed with respect to the more basic domain of PROPERTY. Construal is the way a speaker presents a conceptual representation through the choice of a linguistic expression. While lexemes and constructions bring with them a conventional meaning, this coded meaning is modified by the context in which linguistic units occur. Far from being a fixed, pre-assigned feature, meaning is negotiated locally and socially. By adopting a specific focal adjustment, the speaker linguistically organizes a scene and influences the way the conceptual representation that the expression evokes will be received by the hearer.

This is consistent with the theoretical framework of Cognitive Construction Grammar (CCxG), as laid out in Langacker (1987), Goldberg (1995), and further refined in Goldberg (2003, 2006, 2009) and Langacker (2008, 2009). One of the tenets of CCxG concerns the link between constructions and construal.⁵ Most Construction Grammar supporters recognize that a given construction is both a product and a vector of conceptualization (see Note 1). Goldberg (2003:219) notes that a linguistic pattern counts as a construction “as long as some aspect of its form or function is not strictly predictable from its component parts or from other constructions recognized to exist”. Productive or semi-productive constructions such as *day after day*, *twistin'*

5. The fact that the noun *construction* has two corresponding verbs, *construct* and *construe*, is not innocent.

the night away, or boy, was she in trouble! are idiosyncratic and non-compositional. They must be learned on the basis of the input. Likewise, it may well be the case that entrenched types of the <moderator + adjective> construction provide access to conceptual structures in ways that are not necessarily compositional.

Competence-based versions of Construction Grammar might deny the <moderator + adjective> pattern the constructional status or at least warn that we should not treat *de facto* all collocation patterns as constructions. In Fillmorean Construction Grammar, for example, only collocation patterns that are truly productive count as constructions, other kinds of collocation patterns being consigned the ‘meta-grammar’ together with other ‘patterns of coining’ (Kay 2013). In contrast, usage-based models of Construction Grammar such as CCxG have a broader understanding of the notion of a construction. In a neo-Saussurean fashion, they consider that all form-meaning pairings count as constructions: “it’s constructions all the way down” (Goldberg 2003). To put it plainly, according to competence-based models of Construction Grammar, the linguist has no *a priori* reason to think that a unit counts as a construction, whereas according to usage-based models, linguists have no *a priori* reason to believe that a unit is *not* a construction. In this paper we adopt a usage-based perspective because it leaves it to empirical analysis to decide whether a unit counts as a construction or not. Our choice is reinforced by the fact that the quantitative apparatus designed for constructions can also handle regular collocation patterns (i.e. patterns that competence-based models consider as not necessarily constructional).

The entrenchment of some instances of the <moderator + adjective> construction explains why moderators scale in different ways depending on the adjective they modify. For example, the scaling force of *pretty* depends on the adjective it modifies. As Athanasiadou points out, “*pretty small* is not very small, but *pretty straight* is very/quite straight” (Athanasiadou 2007: 557). More recently, Goldberg (2006: 45) allowed that “individual patterns that are fully compositional are recorded alongside more traditional linguistic generalizations”. This means that all types of the <moderators + adjective> construction, i.e. entrenched as well as non-entrenched types, are worth investigating because each is the trace and the vector of construal mechanisms in the domain of degree modification.

Since corpus-based linguistics provides a comprehensive array of methods to capture context and knowledge, it is not surprising that it has become central in the investigation of cognitive patterns of usage in Cognitive Linguistics (Gries & Stefanowitsch 2006). Paradis (1997: 62) claims that the semantic relation between degree modifiers and adjectives is bidirectional. More precisely, the adjective selects a degree modifier, which in turn restricts the interpretation of the adjective. This means two things. First, we should not examine probabilistic co-occurrences of words regardless of their morphosyntactic environment. Instead, we should inspect co-occurrences of moderators and adjectives within constructional patterns because both moderators and adjectives contribute to the meaning of the <moderator + adjective>

construction. Second, the semantic interaction between a moderator and an adjective is not necessarily compositional.

2.3 Cognitive synonyms or near-synonyms?

According to Paradis (1994:160; 1997:71), moderators are ‘cognitive synonyms’ because substituting one for the other does not change the truth-value of the proposition (Cruse 1986:270ff.).⁶ Cognitive synonyms share identical conceptual content and differ only in style (*die* vs. *pass away*), register (*drunk* vs. *pissed*), and connotation (*firm* vs. *stubborn*). In other words, cognitive synonymy is a matter of both sameness and difference.

Near-synonyms are sometimes referred to as ‘plesionyms’ (Hirst 1995; Edmonds & Hirst 2002; Storjohann 2009). Plesionyms differ from cognitive synonyms because the former involve a slight change in denotational reference, be it in terms of degree (*drunk* vs. *hammered*), fuzzy boundary (*forest* vs. *woods*), viewpoint (*slim* vs. *skinny*), intensity (*break* vs. *destroy*), etc. Substituting near-synonyms alters truth conditions but the sentences where they appear remain semantically similar. Cognitive synonymy and near-synonymy are often hard to tell apart, and an opposition between these two concepts is somehow unhelpful. According to Edmonds and Hirst, near-synonyms bring with them a finer-grained representation than cognitive synonyms. Their conclusion is that “it should not be necessary to make a formal distinction between cognitive synonyms and plesionyms” (Edmonds & Hirst 2002).

The conceptual content that moderators share is essentially functional: *rather*, *quite*, *fairly*, and *pretty* moderate the qualities denoted by the adjectives they modify. But moderators are not completely interchangeable in all contexts. Two reasons have been put forth. First, moderators do not express the same degree of moderation. Second, they involve distinct modes of construal (Paradis 2000, 2008). I will put forth another reason: moderators do not always operate within the same conceptual domains.

3. Method

I combine two broad types of statistics: analytical statistics and multifactorial methods. Analytical statistics uses measures such as frequencies in the domain of hypothesis testing. Rather than formulate and test hypotheses, multifactorial methods aim to formulate a statistical model, i.e. the statistical description of variation and complex relations in a multi-variable dataset.

6. Cruse’s definition of cognitive synonymy echoes Quine’s (1951).

To describe and interpret the distribution of the four moderators, we shall start with a study of collocation patterns. The information concerning the distribution of moderators will serve as input for two methods of multifactorial analysis. The paragraphs that follow will justify the choice of the Corpus of Contemporary American English, explain how the data was extracted, and present the measures of lexico-grammatical co-occurrence that are used, namely collexeme analysis, and multiple distinctive collexeme analysis. Finally, two multifactorial analyses will be introduced: hierarchical cluster analysis and correspondence analysis. Their input will be provided by the results of the collexeme and multiple distinctive collexeme analyses.

3.1 The corpus

At the time of writing, the Corpus of Contemporary American English (Davies, 1990-present), henceforth COCA, consists of 414,771,808 words of spoken and written American English divided among 169,140 texts. The spoken part contains approximately 85 million words and consists of transcripts of conversation from TV and radio programs. The written part is divided evenly between four genres: fiction (80 million words), popular magazines (86 million words), newspapers (82 million words), and academic journals (82 million words). The whole corpus spreads across the period from 1990 to 2010, and 20 million words are added each year. One advantage of COCA is that it is probably the largest publicly and freely available annotated corpus of English. Its size and sampling scheme increase the reliability and validity of observations of relatively rare linguistic phenomena. However, one major disadvantage with COCA is that the texts themselves are not available for download for copyright reasons. The only way to run queries is via the native search interface.⁷ Nevertheless, it is simple to copy and export query results into a text editor or a spreadsheet, clean up the data, cross-tabulate, and run statistical tests.

Davies claims that COCA is balanced. To some extent this is true because words are evenly distributed across genres. Yet, as is the case with all major corpora, balance is more an ideal than a reality. COCA also suffers from the fact that speaker-related metadata are not computed relative to the overall size of the corpus, making them very difficult to measure statistically. For example, COCA indicates the identity of the speakers, but gender is left implicit. However, for the current purposes, the fact that speaker-related information is hard to obtain is of little importance since the main

7. This could be a major issue in light of Leech's "standards of good practice" for corpus users and corpus compilers (Leech 1997:6). At the top of Leech's list is the following recommendation: "[t]he raw corpus should be recoverable." Even if the whole texts are unavailable to the COCA end-users, it is nevertheless possible to "dispense with the annotations, and to revert to the raw corpus" (*ibid.*), although partially.

focus of the present study is on function-related distributions and collocation patterns, regardless of dialectal issues.

Last but not least, queries on the COCA yield duplicates. This is largely due to automatic text collection, a small price to pay for a corpus of this size. I have not removed duplicates for two reasons. First, duplicates that result from a genuine verbatim quotation are very hard to distinguish from duplicates that result from an error in automatic text compilation. Second, the number of duplicates does not affect the statistics in any significant way. All corpora come with restrictions, and it is important to bear these in mind when interpreting results. All things considered, the advantages of COCA outnumber its disadvantages.

3.2 From collocates to collexemes

A common assumption in corpus linguistics is that the context of a variable (lexical or phrasal) reveals important aspects of its syntactic and semantic properties (Sinclair 1991; Biber *et al.* 1998). The easiest way to analyze the context of a variable is to extract its collocates and determine those that most typically combine with the variable. According to Sinclair (1991: 170), collocation is “the occurrence of two or more words within a short space of each other in a text”. Traditionally, the conventional size of what Sinclair calls ‘a short space’ varies from 1 to 5 words on both sides of the variable, depending on the case study. Syntactic studies tend to examine wider spans, while lexical studies (e.g. those that focus on idiomatic compounds such as *bread and butter*) examine shorter spans. Some authors also assume that statistical methods can easily filter out the noise generated by a wide concordance span, but once again the wider the span, the weaker the syntactic claims the linguist can make.

In the literature on intensifiers, little is said as to the optimal span to investigate. In his study on amplifiers, Kennedy inspects “a window of two words” on each side so as to “retrieve collocations that may have been separated by intervening words” (2003: 472–473). Amplifiers can modify elements that precede or follow. This fact alone justifies a two-way search. Amplifier collocates can be adjectives, other adverbs, or verbs, and given their multifunctionality, “intervening words” are frequent. Moderators are slightly different. Insofar as they are ‘word modifiers’ (Stoffel 1901), there is no need to inspect too wide a range of words, and since these adverbs are premodifiers it is useless to inspect the left context.⁸ Bearing these specificities in mind, I extracted

8. In many collocation-based studies, node words or phrases are often clustered on the basis of their collocates within a relatively large span. One big problem with this method is that the semantic relation between many of these collocates and the node(s) is loose, which brings about noise. By adopting a constructional approach, and therefore treating the sequence <moderator + adjective> as a construction, I restrict the semantic investigation to the syntactic frame of the construction and inevitably minimize the risk of obtaining noise in the data points.

all adjectives that occur in the first two slots to the right of *rather*, *quite*, *fairly*, and *pretty*. I also adopted the “paradigmatic reduction” outlined in Lorenz (2002: 144) and considered only moderators of adjectives so as to focus on subtle semantic relations.

Collocation extraction is a preliminary step to determining sequences of moderators and adjectives that co-occur more frequently than would be expected by chance. Determining collocation strength is a particularly helpful way of spotting differences between semantically similar expressions. Even though we assume that *rather*, *quite*, *pretty* and *fairly* are near-synonyms, and therefore map onto similar content domains in a quality-related conceptual space, we may expect significantly distinct collocation patterns for each moderator. A related assumption is that distinct collocation patterns reflect subtle differences in schematic-domain profiling.

As mentioned earlier, the semantic relation between degree modifiers and adjectives has been described as bidirectional. While this is a convincing assumption, a simple collocate-based inquiry based on raw counts and/or basic relative frequencies is unsatisfactory because it fails to distinguish those adjectives that are significantly attracted to a degree modifier from those that are frequent in the corpus regardless of the context where they appear.⁹

In this study of the <moderator + adjective> construction, we shall turn to collexeme analysis (Stefanowitsch & Gries 2003), a method that can both handle two-way semantic influences and filter out adjectives that have a high overall token frequency in the corpus. Collexeme analysis is part of a family of methods known as collostructional analysis (Hilpert, this volume, 391–404).¹⁰ It investigates which lexemes typically occur in a given slot in a single grammatical construction (e.g. *the X-er, the better*). It takes as input the frequency of a lexeme in a given construction, the frequency of the same lexeme in all other constructions, the frequency of the construction with other lexemes, and the frequency of all other constructions with other lexemes (Stefanowitsch & Gries 2003: 218). The data is tabulated and the 2x2 table is submitted to the Fisher-Yates Exact test (Pedersen 1996). Unlike χ^2 -statistics, the Fisher-Yates Exact test does not presuppose or violate any distributional assumptions (Kilgarriff 2005). Unlike χ^2 -statistics and Mutual Information, the Fisher-Yates Exact test is not invalid when counts are low. This means that rare collocations in COCA (such as *quite medical* or *pretty social*) can nevertheless be included in the overall calculation of collostruction strength. The *p*-value that one obtains through the Fisher-Yates Exact test is an indication of the association strength between the lexeme and the construction:

9. Raw counts provide the simplest association measure for pair types. When the construction under investigation displays low token frequency, raw counts are often the only option. However, they should be regarded as a last resort.

10. Each of these methods measures the degree of attraction or repulsion between lexical items and constructions. In collostructional analysis, a construction is defined as a conventional pairing of form and meaning, in light of Goldberg (1995).

the smaller the p -value, the stronger the association between a lexeme and a construction. In recent versions of collexeme analysis, the p -value is transformed in an inverse logarithmic function so as to make the distinction between very small p -values easier to identify. In the present study, this operation will be repeated for all adjective types that co-occur with each moderator. Collexeme analysis will first help determine which adjectives are most strongly attracted to the construction <moderator + adjective> by quantifying the bidirectional attraction between a moderator and the adjective it modifies. Second, it will help confirm that *rather*, *quite*, *pretty*, and *fairly* are functional synonyms by revealing significant overlap in the selection of adjectives. The more significant the overlap, the more these four moderators map onto similar content domains, and the stronger the claim that moderators are cognitive synonyms.

Cognitive synonyms also display differences. To make these differences apparent, one can look for the distinctive collexemes of each construction. This is where a second method of collostructional analysis is needed: distinctive collexeme analysis (Gries & Stefanowitsch 2004). Insofar as we want to compare four near-synonymous constructions, <*rather* + adjective>, <*quite* + adjective>, <*pretty* + adjective>, and <*fairly* + adjective>, and contrast them in their respective collexeme preferences, we should implement multiple distinctive collexeme analysis (Hilpert 2006; Gilquin 2007). This method differs from simple collexeme analysis because it filters away overlapping collocates and retains only those adjectives that are specific to each moderator. The output makes it possible to classify the distinctive adjectives according to their function and meaning, and to get a better grasp of the individual functional specificities of *rather*, *quite*, *pretty*, and *fairly*.

3.3 Collexemes as input for multivariate statistics

Moderator collexemes can serve as input for two usage-based techniques whose aim is to capture semantic relations between near-synonyms on the basis of multiple factors: hierarchical cluster analysis and correspondence analysis. Both methods are exploratory. Instead of testing a hypothesis in relation to pre-assigned categorical clusters, multifactorial analyses follow a bottom-up approach. Indeed, clusters are determined by the similarity of the members of the same groupings and their dissimilarity to the members of other groupings. In theory, these methods dispense with the linguist's preconception of how the data is categorized. But in practice the clustering results in large part from the criteria that the linguist adopts to combine points into clusters.

Recent research on lexical near-synonymy in Cognitive Linguistics has made extensive use of cluster analysis (Divjak 2006, 2010; Divjak & Fieller this volume, 405–441; Divjak & Gries 2008). Hierarchical cluster analysis is the generic name of a family of statistical techniques for clustering data (i.e. for structuring observed data into groups) and representing them graphically in a tree-like format. Among these clustering techniques, I use hierarchical agglomerative clustering, whereby individual

data points are successively agglomerated into similar clusters, and similar clusters are merged iteratively into bigger clusters until one last cluster is obtained.¹¹ Results appear in the form of a dendrogram, which facilitates an objective and accurate identification of semantic classes and subclasses for a given lexical set. Hierarchical agglomerative cluster analysis will be used to justify the existence of the paradigm of moderators within the broader paradigm of degree modifiers of adjectives. The collocation strength of the collexemes of 23 English degree modifiers of adjectives will serve as input to compute the distance matrix.

Correspondence analysis is another distance-based clustering technique that represents the structure of cross-tabulations graphically on a multi-dimensional plane (Benzécri 1973; Benzécri 1984; Greenacre 2007; Glynn, this volume, 443–486). In lexical semantics, correspondence analysis offers a convenient way of mapping correlations between lexical items graphically. Like hierarchical cluster analysis, correspondence analysis has become a popular method in Cognitive Semantics (Glynn 2010a). Interpreting a correspondence map is relatively simple: the closer the data points on the map, the stronger the correlation between these data points. However, interpreting a map can also be tricky, since it flattens multi-dimensional distances onto a two-dimensional plane. Nevertheless, correspondence analysis yields promising results in areas such as exemplar-based semantics or cognitive sociolinguistics (Glynn 2010b). Cognitive linguists like the fact that a correspondence map may provide access to the complex conceptual maps that structure language knowledge and language use. Correspondence analysis is used to show how moderators imply a specific construal depending on the adjectives they correlate with. Input is provided by the cross-tabulation of the frequencies of 25 distinctive collexemes for each moderator. The list of 25 distinctive collexemes will be obtained via the multiple distinctive collexeme analysis outlined above.

This paper attempts to show that near-synonymy is a complex phenomenon whose study can only benefit from a combination of different statistical methods. These methods range from frequencies and collocations – including (multiple distinctive) collexeme analysis – to exploratory multifactorial techniques – namely cluster analysis and correspondence analysis. If cognitive synonymy is indeed a matter of sameness and difference, only a combination of statistical techniques can provide a sound-enough basis for a complete picture of semantic relations to emerge regarding moderators.

11. Hierarchical divisive clustering proceeds in reverse order: it starts at the root and successively splits the clusters.

4. Results

4.1 Collexeme analysis

Because *rather*, *quite*, *pretty*, and *fairly* are alternate ways of expressing moderation, we should expect them to display similarities and differences. Each feature will be illustrated in turn, starting with similarities, which will be captured by means of a collexeme analysis. Collexeme analysis generates a ranked list of attracted lexemes (i.e. ‘collexemes’) and rejected lexemes. This list can be used to determine what meanings are congruent with the semantics of the construction and what meanings are incongruent. For our current purpose, we shall focus on congruent meanings.

Table 1. Top 10 adjectival collocates of *rather*, *quite*, *fairly*, and *pretty* in COCA

<i>rather</i>			<i>quite</i>		
adjective	frequency in corpus	frequency in construction	adjective	frequency in corpus	frequency in construction
<i>large</i>	119,992	260	<i>different</i>	17,021	2,247
<i>different</i>	17,021	231	<i>sure</i>	137,372	1,347
<i>small</i>	165,348	189	<i>clear</i>	81,553	805
<i>difficult</i>	6,672	129	<i>good</i>	378,826	578
<i>unusual</i>	17,736	116	<i>right</i>	4,558	548
<i>simple</i>	48,134	96	<i>possible</i>	88,919	458
<i>limited</i>	3,533	95	<i>similar</i>	60,967	328
<i>good</i>	378,826	89	<i>ready</i>	55,338	300
<i>high</i>	191,591	85	<i>common</i>	63,239	278
<i>strange</i>	23,457	80	<i>simple</i>	48,134	271
<i>fairly</i>			<i>pretty</i>		
adjective	frequency in corpus	frequency in construction	adjective	frequency in corpus	frequency in construction
<i>good</i>	378,826	346	<i>good</i>	378,826	7,492
<i>easy</i>	59,914	337	<i>sure</i>	137,372	1,241
<i>large</i>	119,992	281	<i>bad</i>	90,297	743
<i>common</i>	63,239	278	<i>clear</i>	81,553	729
<i>high</i>	191,591	247	<i>big</i>	187,641	583
<i>simple</i>	48,134	243	<i>tough</i>	33,746	486
<i>new</i>	64,824	202	<i>cool</i>	3,235	481
<i>certain</i>	71,149	201	<i>close</i>	94,845	471
<i>small</i>	165,348	190	<i>hard</i>	123,894	436
<i>typical</i>	20,483	154	<i>strong</i>	69,137	383

Let us postulate that each moderator of adjectives forms a construction that attracts certain adjectival collexemes and rejects others. To calculate the collocation strength of a given adjective *A* for a given <moderator + adjective> construction *C*, collexeme analysis needs four frequencies: the raw frequency of *A* in *C*, the raw frequency of *A* in all other constructions (i.e. $\neg C$), the frequency of *C* with adjectives other than *A* (i.e. $\neg A$), and the frequency of $\neg C$ with that of $\neg A$ (Stefanowitsch & Gries 2003:218, Hilpert this volume). To generate a ranked list of attracted adjectives based on collocation strength, this operation is repeated for each adjective that co-occurs with each moderator in the corpus. Table 1, above, summarizes the frequencies of the 10 most frequent adjectives in the four <moderator + adjective> constructions in COCA, i.e. <*rather* + adjective>, <*quite* + adjective>, <*fairly* + adjective> and <*pretty* + adjective>.

Each of the four subcomponents of the table are submitted to a collexeme analysis, along with the following information:

corpus size: 415M words
 frequency of <*rather* + adj.>: 12,574
 frequency of <*quite* + adj.>: 29,735
 frequency of <*fairly* + adj.>: 10,834
 frequency of <*pretty* + adj.>: 35,949

Table 2 presents the ten most strongly attracted adjectives of the four moderators based on collocation strength.¹²

Table 2. Top 10 collexemes of *rather*, *quite*, *fairly*, and *pretty*

<i>rather</i>		<i>quite</i>		<i>fairly</i>		<i>pretty</i>	
adjective	coll. strength	adjective	coll. strength	adjective	coll. strength	adjective	coll. strength
<i>large</i>	1025.19	<i>different</i>	15082.43	<i>easy</i>	2686.55	<i>good</i>	61820.34
<i>different</i>	709.26	<i>sure</i>	8231.32	<i>common</i>	2078.88	<i>sure</i>	8148.39
<i>unusual</i>	705.49	<i>clear</i>	4923.96	<i>simple</i>	1883.49	<i>clear</i>	4764.46
<i>small</i>	521.34	<i>possible</i>	2216.18	<i>large</i>	1751.18	<i>bad</i>	4734.36
<i>difficult</i>	480.6	<i>similar</i>	1614.27	<i>good</i>	1523.3	<i>tough</i>	3635.13
<i>odd</i>	436.5	<i>good</i>	1482.02	<i>straight-forward</i>	1433.43	<i>cool</i>	3628.34
<i>remarkable</i>	415.88	<i>ready</i>	1480.81	<i>certain</i>	1326.03	<i>amazing</i>	2848.64
<i>limited</i>	413.37	<i>simple</i>	1357.46	<i>typical</i>	1315.09	<i>big</i>	2604.22
<i>vague</i>	400.68	<i>remarkable</i>	1297.76	<i>high</i>	1250.31	<i>close</i>	2531.19
<i>strange</i>	384.69	<i>common</i>	1259.48	<i>consistent</i>	1112.44	<i>strong</i>	2139.59

12. It was performed thanks to the script Coll.analysis 3.2 for R (Gries 2007).

One might object that there is little difference between the ranked list based on raw frequencies (Table 1) and the ranked list based on collocation strength (Table 2). However, the latter is more reliable since collocation strength is the product of an association measure – here Dunning’s log-likelihood ratio – which uses absolute and collocate frequencies to determine which lexemes co-occur more frequently than expected in a given construction. All the above collexemes are attracted to each construction at the very significant level of $p < 0.001$ since $\text{coll.strength} > 3$.¹³

Upon inspection, collexeme analysis informs a semantic analysis of moderator constructions. The collexemes of *rather* split into sets of adjectives that denote spatial dimension (*large, small, limited*), atypicality (*unusual, odd, vague, strange, remarkable*), difference, and difficulty. The collexemes of *quite* split into the following classes: difference/similarity (*different, similar*), epistemic modality (*sure, possible*), dynamic modality (*ready*), positive value (*good, clear*), (a)typicality (*common* vs. *remarkable*), and simplicity. As regards *fairly*, collexemes divide up into adjectives that denote simplicity (*easy, simple, straightforward*), typicality (*common, typical, consistent*), spatial dimension or position (*large, high*), positive value (*good*), and epistemic modality (*certain*). The collexemes of *pretty* fall into the following sets: positive/negative values (*good* vs. *bad, cool, clear, strong*), difficulty (*tough*), nonstandard identification (*amazing*), spatial dimension or position (*big, close*), and epistemic modality (*sure*).

As expected, collexeme analysis reveals that the behavioral profiles of moderators are both similar and different. Similarity is evidenced by the significant degree of overlap in the selection of collexemes. Many adjectives co-occur with more than one moderator (*certain, clear, common, different, difficult, good, large, remarkable, simple, sure*). Some semantic classes are shared among moderators. For example, both *rather* and *quite* are compatible with the expression of (a)typicality. Moderating qualities that belong to the class of spatial extension can be done by means of *fairly* or *rather*. The expression of epistemic modality is a feature common to *quite, pretty, and fairly*. This confirms that moderators are synonyms to some extent.

However, at a finer-grained level of analysis, semantic correspondences between moderators are not that straightforward. Even though the expression of atypicality is observed with *quite*, it is in fact a distinctive feature of *rather* because adjectives denoting that semantic class are more varied with *rather*. Even though the expression of epistemic modality is common to the three moderators, it is actually characteristic of *quite*. The adjective *different* co-occurs with both *rather* and *quite*. However, a quick look at the collocation strength shows that *different* is far more distinctive of *quite* ($\text{coll. strength} = 15082.43$) than it is of *rather* ($\text{coll. strength} = 709.26$). *Good* is attracted to *pretty, fairly, and quite*, but it is definitely more distinctive of *pretty*

13. As a rule, if collocation strength is based on p -values, the following equivalences hold true: $\text{coll.strength} > 3 = p < 0.001$; $\text{coll.strength} > 2 = p < 0.01$; $\text{coll.strength} > 1.3 = p < 0.05$.

(coll. strength = 61820.34) than it is of *fairly* (coll. strength = 1523.3), and *quite* (coll. strength = 1482.02).¹⁴

Collexeme analysis is not the easiest way of spotting differences in collocational preferences because it does not filter away overlapping adjectives and it requires a tedious comparison of collocation strengths to determine relevant thresholds of attraction. This can be done by means of a multiple distinctive collexeme analysis (MDCA), which contrasts constructions in their distinctive collocational preferences by getting rid of overlapping collexemes. As explained above, this method is best suited for comparing related constructions, preferably alternations. Whether *rather*, *quite*, *pretty*, and *fairly* are alternative ways of expressing moderation is undeniable. However, the internal structure of the paradigm that these four adverbs belong to is problematic because some of these adverbs (e.g. *quite*) can be used as maximizers. To make sure the paradigm of moderators is internally coherent and thus to maximize the interpretation of MDCA, one interesting option is to conduct a hierarchical cluster analysis over a pool of 23 degree modifiers in English and see if the four adverbs cluster together on the basis of their preferred collexemes.

4.2 Collexeme analysis as input for hierarchical cluster analysis

Hierarchical cluster analysis describes a range of multifactorial methods for investigating structure in data, with the goal of identifying subgroups of similar objects. Following Gries & Stefanowitsch (2010), I use hierarchical agglomerative clustering (Everitt *et al.* 2011: Section 4.2) to see how English degree modifiers cluster on the basis of their preferred collexemes. The 23 degree modifiers, which include the four moderators under investigation, are: *a bit*, *a little*, *absolutely*, *almost*, *awfully*, *completely*, *entirely*, *extremely*, *fairly*, *frightfully*, *highly*, *jolly*, *most*, *perfectly*, *pretty*, *quite*, *rather*, *slightly*, *somewhat*, *terribly*, *totally*, *utterly*, *very*. Originally, Paradis selected them because they epitomize the degree modifier paradigm in most lexicographic works (1997: 15–17). If *rather*, *quite*, *pretty*, and *fairly* cluster together, then these four adverbs form a homogeneous paradigm despite their multifunctional behavior.

For each of the 23 adverb types listed above, I first extracted all adjectival collocates in COCA, amounting to 432 adjective types and 316,159 co-occurrence tokens. Then, I conducted a collexeme analysis for each of the 23 degree modifiers. To reduce the data set to manageable proportions, the 35 most attracted adjectives were selected on the basis of their collocation strength. For these 23 adverb types and their 432 adjective types, a 23-by-432 co-occurrence table containing the frequency of each adverb-adjective pair type was submitted to a hierarchical agglomerative cluster analysis, which requires a distance object as input. The distance object is a dissimilarity matrix

14. It is also a collexeme of *rather*, although to a much lesser extent (coll. strength = 34.34).

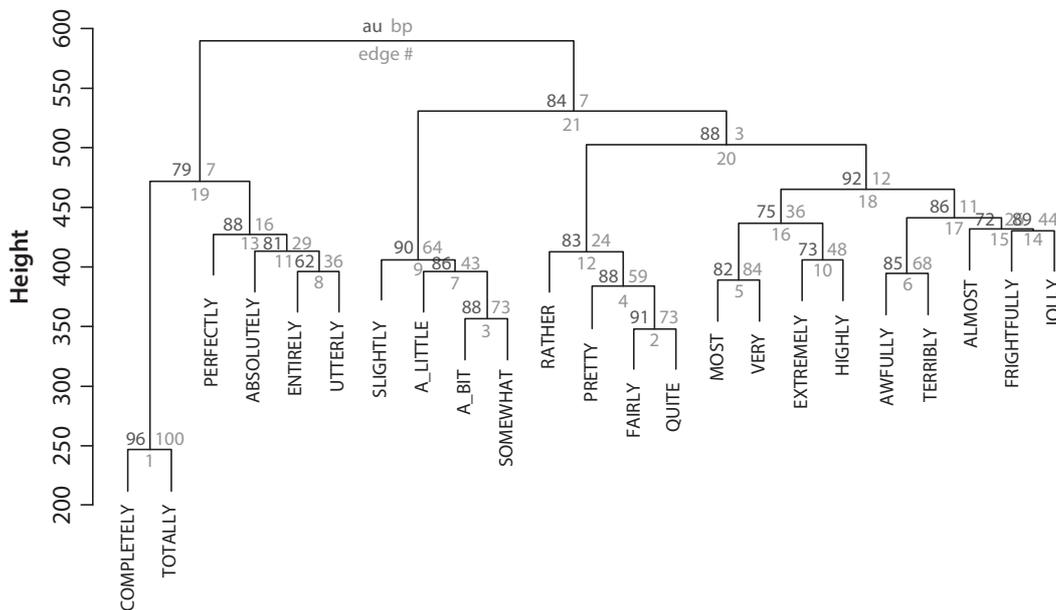


Figure 1. Cluster dendrogram of 23 degree modifiers of adjectives in English, clustered according to their adjectival collexemes (distance: Canberra; cluster method: Ward)

that one obtains by converting tabulated frequencies into distances with a user-defined distance measure. When variables are ratio-scaled, the linguist can choose from several distance measures (Euclidean, City-Block, correlation, Pearson, Canberra, etc.).¹⁵ For our purpose, the measure of dissimilarity of the adverb types in the columns was computed using the Canberra distance metric, because it handles the relatively large number of empty occurrences best (see Divjak & Gries [2006: 37] for further methodological details). Finally, one needs to apply an amalgamation rule that specifies how the elements in the distance matrix get assembled into clusters. Here, clusters were amalgamated using Ward's method (Ward 1963), which evaluates the distances between clusters using an analysis of variance. This method has the advantage of generating clusters of moderate size.¹⁶ Figure 1 shows the resulting dendrogram.

The plot should be read from bottom to top. There are three numbers around each node. The number below each node specifies the rank of the cluster (here, from 1 to 21, i.e. from the 1st generated cluster to the 21st). The two numbers above each node

15. For reasons of space, I cannot discuss the reasons why one should prefer a distance measure over another. A description of some distance measures can be found in Gries (2010: 313–316).

16. All computations were performed with R 2.13 (R Development Core Team 2011) and the package `pvclust` (version 1.2-2, www.is.titech.ac.jp/~shimo/prog/pvclust/). This package allows the user to include confidence estimates through multiscale bootstrap resampling, a possibility missing in other packages, such as `hclust`.

indicate two types of p -values,¹⁷ which are calculated via two different bootstrapping algorithms: AU and BP. The number to the left indicates an ‘approximately unbiased’ p -value (AU) and is computed by multiscale bootstrap resampling. The number to the right indicates a ‘bootstrap probability’ p -value (BP) and is computed by normal bootstrap resampling. The number to the left is a much better assessment of how strongly the cluster is supported by the data. In both cases, the closer the number is to 100, the stronger the cluster. AU p -values suggest the clusters we obtain represent the data accurately. Indeed, the plot shows the standard values of most clusters are significantly high, with AU p -values ranging from 79 to 96. An AU p -value of 96 implies that the hypothesis that the cluster is invalid is rejected with a significance level of 0.04.

The dendrogram displays several homogeneous clusters:¹⁸

- a. cluster 19 groups together maximizers; it breaks down into cluster 1 (*completely, totally*) and cluster 13 (*perfectly, absolutely, entirely, utterly*);
- b. cluster 9 groups together diminishers (*slightly, a little, a bit, somewhat*);
- c. cluster 12 groups together moderators (*rather, pretty, fairly, quite*);
- d. cluster 18 groups together boosters and breaks down into cluster 16 (*most, very, extremely, highly*), cluster 6 (*awfully, terribly*), and cluster 14 (*frightfully, jolly*); the presence of an approximator (*almost*) within the cluster of boosters (cluster 15) is surprising but may be due to its intensive use as a sentential adverb, more than a modifier of adjectives.¹⁹

The cluster analysis based on collexemes yields functionally and semantically motivated groups. As Paradis (1997:27) observed, *rather, quite, pretty, and fairly* do cluster together under the moderator paradigm (cluster 12) despite their multifunctionality. However, the internal structure of this cluster still needs explaining. It is not clear why *fairly* and *quite* cluster together (cluster 2), and why *rather* is not part of cluster 4, which groups together *pretty* and cluster 2. Furthermore, the stratification of cluster 12 does not follow their conventional distribution in terms of grading force (*rather* > *quite* > *pretty* > *fairly*), as found in Paradis (1997: 148–155). For now, suffice it is to say that moderators form a functionally coherent class. Performing MDCA to explore internal distinctions is therefore relevant.

17. The term “ p -value” is the one that the authors of the `pvclust` package have adopted. Actually, it seems that these p -values are confidence estimates.

18. In the classification of degree modifiers that follows, I adopt Paradis’s terminology.

19. See Paradis (1997:37) for confirmation.

4.3 Multiple distinctive collexeme analysis

We saw above that *rather*, *quite*, *pretty*, and *fairly* display similarities and differences. One way to amplify these differences is to conduct a multiple distinctive collexeme analysis. Instead of computing the degree of attraction between a lexical item and a construction, distinctive collexeme analysis contrasts constructions in their respective collocational preferences (Gries & Stefanowitsch 2004). This method has proved useful when it comes to distinguishing minimal semantic and functional differences between near-synonymous constructions (e.g. the ditransitive *vs* prepositional dative alternation). The input is slightly different from what we have in collexeme analysis. This time, one needs to tabulate the type frequency of the collexeme in the first construction, the type frequency of the same collexeme in the second construction, and the frequencies of the two constructions with words other than the collexeme under investigation (Gries & Stefanowitsch 2004: 102). Again, 2x2 tables are submitted to the Fisher-Yates Exact test for each relevant lexeme. However, when one wants to compare more than two constructions and input more complex tables, such as Table 3 below, the Fisher-Yates Exact test cannot be used. Instead, one needs to carry out a one-tailed exact binomial test, and the method goes under the name of multiple distinctive collexeme analysis.²⁰ The same script as the one used for collexeme analysis was used.

Below, Tables 4 to 7 list, for each moderator, the ten most distinctive adjectives. MDCA compares the observed frequency of each adjective with its expected frequency.

If adjectives were distributed at random over the different moderator constructions, we would not find any significant deviation between observed and expected frequencies because the distribution of each adjective would follow the frequencies of the moderators. For each construction token, the script performs a binomial test

Table 3. Input for a multiple distinctive collexeme analysis of an adjective in four moderator constructions

Construction	Adjective A	Other adjectives	Row totals
<i>rather</i> + adj	a	b	a+b
<i>quite</i> + adj	c	d	c+d
<i>pretty</i> + adj	e	f	e+f
<i>fairly</i> + adj	g	h	g+h
column totals	a+c+e+g	b+d+f+h	a+b+c+d+e+f+g+h

20. Gilquin (2007) illustrates how multiple distinctive collexeme analysis determines the verbs that are distinctively associated with the non-finite verb slot of English periphrastic causative constructions.

Table 4. The 10 most distinctive adjectives of *rather*

<i>rather</i> + adj	observed frequency	expected frequency	pbin <i>rather</i>	SumAbsDev
<i>odd</i>	74	14.51	38.99	56.75
<i>unusual</i>	116	37.58	30.40	46.39
<i>strange</i>	80	25.70	21.54	38.70
<i>vague</i>	56	12.71	24.75	35.86
<i>difficult</i>	129	65.35	13.89	35.86
<i>simplistic</i>	28	4.70	18.31	28.42
<i>lengthy</i>	36	9.26	13.81	27.53
<i>peculiar</i>	31	5.94	17.21	27.08
<i>bizarre</i>	51	14.09	17.46	25.83
<i>curious</i>	29	5.94	14.91	25.09
<i>formal</i>	29	5.94	14.91	23.49

Table 5. The 10 most distinctive adjectives of *quite*

<i>quite</i> + adj	obs freq	exp freq	pbin <i>quite</i>	SumAbsDev
<i>different</i>	2247	825.30	Inf	Inf
<i>right</i>	548	184.75	238.72	410.01
<i>possible</i>	458	150.95	216.96	374.49
<i>ready</i>	300	104.02	120.33	206.00
<i>true</i>	235	91.23	70.29	120.24
<i>likely</i>	206	76.46	69.12	118.13
<i>similar</i>	328	148.32	66.10	116.57
<i>capable</i>	152	51.19	66.87	114.97
<i>willing</i>	142	49.55	56.31	96.25
<i>correct</i>	106	36.10	45.22	78.33

Table 6. The 10 most distinctive adjectives of *fairly*

<i>fairly</i> + adj	obs freq	exp freq	pbin <i>fairly</i>	SumAbsDev
<i>common</i>	278	83.86	76.85	145.26
<i>easy</i>	337	108.56	83.94	130.39
<i>new</i>	202	43.38	88.58	123.43
<i>constant</i>	117	16.63	84.16	123.41
<i>recent</i>	122	20.36	72.50	109.92
<i>certain</i>	201	66.03	49.00	98.80
<i>typical</i>	154	35.55	61.99	89.24
<i>consistent</i>	130	33.26	45.98	68.47
<i>straightforward</i>	123	30.73	44.94	65.25
<i>regular</i>	70	12.05	40.51	59.78

Table 7. The 10 most distinctive adjectives of *pretty*

<i>pretty</i> + adj	obs freq	exp freq	pbin <i>pretty</i>	SumAbsDev
<i>good</i>	7731	3613.04	Inf	Inf
<i>bad</i>	758	343.37	201.75	349.84
<i>cool</i>	488	214.45	144.61	252.23
<i>tough</i>	489	226.02	122.02	213.70
<i>big</i>	591	295.44	113.65	204.40
<i>hard</i>	447	225.61	83.52	143.97
<i>scary</i>	212	98.34	52.78	93.43
<i>smart</i>	196	91.32	48.17	83.66
<i>amazing</i>	347	195.03	44.89	82.05
<i>close</i>	498	314.03	40.52	77.60

to determine the probability of a particular observed frequency given the expected frequency.²¹ This probability is then log-transformed. The resulting value (pbin) captures distinctiveness.²² It is used to determine whether a given adjective is distinctive for a particular construction or not, and whether the co-occurrence between the adjective and the moderator construction is statistically significant or not. The co-occurrence is statistically significant if the absolute distinctiveness value is higher than 1.3, $p < 0.05$. Finally, SumAbsDev gives the sum of all absolute pbin values for a particular adjective. The higher the figure, the more the adjective deviates from its expected distribution.

MDCA makes patterns of attraction more visible. It confirms that *rather* attracts adjectives that denote atypicality/deviation from a norm (*odd, unusual, strange, vague, peculiar, bizarre, curious*). Additionally, *rather* attracts adjectives that denote difficulty/simplicity. By far, the most distinctive collexeme of *quite* is *different* (pbin and SumAbsDev = infinite). Its antonym (*similar*) is also among the 10 most distinctive collexemes. Modal meanings are well represented: *possible* and *likely* denote epistemic meaning, and *ready, capable, and willing* denote dynamic meaning. Also distinctive of *quite* are adjectives that denote factuality (*right, true, correct*). *Fairly* attracts some sets that are semantically close, such as typicality (*common, typical*), similarity/stability (*constant, consistent, regular*), and epistemicity (*certain*). Other sets include easiness (*easy, straightforward*), and time location (*new, recent*). Lastly, the most distinctive collexeme of *pretty* is *good* (pbin and SumAbsDev = infinite). *Good* belongs to the

21. For instance, the probability to find 74 occurrences of *odd* in *<rather + adj.>* when you would have expected it 14.51 times.

22. It receives a positive sign when the verb occurs more frequently than expected in the construction and a negative sign when the verb occurs less frequently than expected. In short, positive values indicate attracted collexemes whereas negative values indicate repelled collexemes. For reasons of space, I have selected positive values only.

category of positive values, along with *cool* and *smart*. *Pretty* also attracts an antonym such as *bad*, which denotes a negative value. Other distinctive semantic sets include difficulty (*tough*, *hard*), spatial dimension or location (*big*, *close*), deviation from a norm (*amazing*), and psychological stimulus (*scary*).

Since MDCA excludes overlap (i.e. those adjectives which collexeme analysis revealed as common to at least two moderators), it makes some tendencies that collexeme analysis revealed more apparent:

- a. atypicality as well as difficulty/simplicity are the most distinctive features of *rather*;
- b. difference/similarity and modal meanings are the most distinctive features of *quite*;
- c. typicality is a distinctive feature of *fairly*, along with similarity/stability;
- d. whatever their polarity, value judgments are the most distinctive features of *pretty*, along with difficulty and dimension/position.

But MDCA also reveals tendencies that were harder to grasp with collexeme analysis. Indeed, moderators follow a division of labor in the expression of some functions. Atypicality is a distinctive feature of *rather*, whereas typicality is a distinctive feature of *fairly*. Easiness is a distinctive feature of *fairly*, whereas the expression of difficulty is distinctive of both *rather* and *pretty*. There is a difference in register though: it seems that *pretty* is less formal than *rather* (*rather difficult* vs. *pretty tough*, *pretty hard*). All in all, MDCA shows that even though moderators are functionally close, they do not profile the same conceptual domains.

Even though collexeme analysis and MDCA reveal tendencies that were much harder to capture with only raw frequencies, the above observations are partial because of the limited number of selected collexemes. For a deeper assessment of the synonymy of moderators and the division of labor that they follow, we should increase the level of granularity of our analysis. One obvious solution is to investigate more collexemes, but the more data we have, the more difficult it is to make generalizations. Rather than inspect and compare collocation-based frequency tables manually, we should also be able to compute and visualize the relative attraction between (a) moderators, (b) adjectives, (c) moderators *and* adjectives. With this goal in mind, we can use the output of MDCA as input for correspondence analysis.

4.4 Multiple distinctive collexeme analysis as input for correspondence analysis

Correspondence analysis (henceforth CA) is an exploratory statistical technique that takes the frequencies of multiway tables as input, then summarizes and visualizes distances between the variables. It determines the probability of global association

Table 8. Input for correspondence analysis (sampled)

Adjective (distinctive collexeme)	<i>fairly</i>	<i>pretty</i>	<i>quite</i>	<i>rather</i>
<i>able</i>	3	4	67	0
<i>abstract</i>	14	7	5	22
<i>accurate</i>	83	66	109	4
<i>amazing</i>	6	347	97	22
<i>aware</i>	1	11	109	0
<i>awful</i>	0	100	12	6
<i>awkward</i>	0	12	7	27
<i>bad</i>	19	758	32	22
<i>beautiful</i>	1	2	129	21
<i>big</i>	56	591	45	23
...

between rows and columns, and tests this association using the χ^2 test.²³ Two rows/columns will be close to each other if they associate with the columns/rows in the same way.

Table 8 shows a sample of the input used for CA. It brings together the 25 most distinctive collexemes of each moderator and the raw frequency of each collocation type. The whole table contains 400 cells. CA uses these frequencies to compare (a) line profiles, i.e. adjectives, (b) column profiles, i.e. moderators, (c) line profiles and column profiles, i.e. moderators and adjectives. This method reintroduces overlap because the table contains raw frequencies of adjectives that co-occur with at least two moderators. As we saw above, overlap is a characteristic of the paradigm of moderators. Taking overlap into account is therefore a way of mapping co-occurrence patterns more realistically than if we simply ignored it.

CA transposes the multidimensional distances to a two-dimensional plane that maps the correlations between the variables. More precisely, it transforms the input table (i.e. a table of numerical information) into a graphic display in which each row and each column is represented as a point in a Euclidean space. Figure 2, below, is the graphic output of CA.²⁴

23. Since CA is an exploratory technique, one does not need to check whether the conditions of use of χ^2 -statistics are met. For our current purpose, the hypothesis of independence can be rejected because $\chi^2 = 33623.82$, $df = 297$, and p -value $< 2.2e-16$.

24. To conduct CA and output the graph, I used R with the packages FactoMineR (<http://cran.r-project.org/web/packages/FactoMineR/index.html>) and dynGraph (<http://cran.r-project.org/web/packages/dynGraph/index.html>).

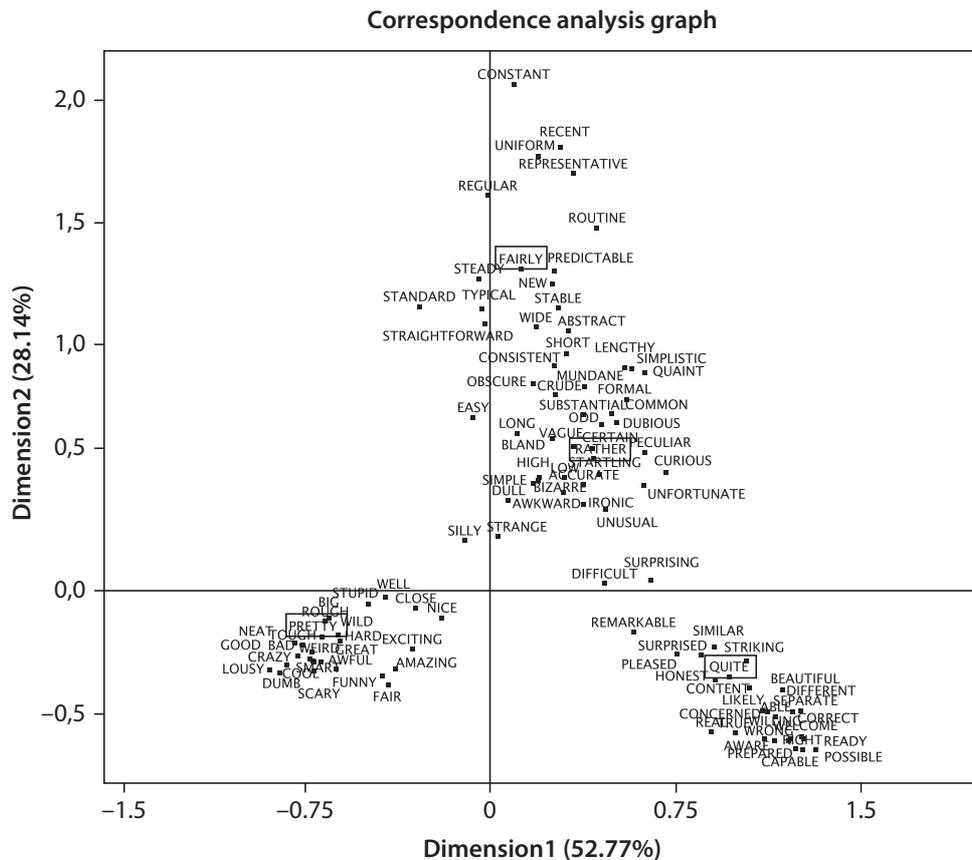


Figure 2. CA biplot of the <moderator + adjective> construction in COCA

The plot is built along two axes, which are the principal axes of inertia.²⁵ Their intersection defines the average profile of all the points in the cloud. CA decomposes the overall inertia by identifying a small number of representative dimensions. Each axis corresponds to a dimension. The plot displays only two dimensions, which are selected according to their eigenvalues. The eigenvalue of a dimension measures how much information is present along the axis of that dimension. The first axis (dimension 1, eigenvalue = 0.533) represents 52.77% of the inertia, whereas the second axis (dimension 2, eigenvalue = 0.284) represents 28.14% of the inertia. There is a third dimension, whose eigenvalue is 0.193.

Even though dimension 3 accounts for 19.08% of the inertia, it is not taken into account in the plot. This is not a problem because the first two dimensions already explain 80.91% of the information contained in the input table, and the results can be interpreted with enough accuracy without dimension 3.

Because the plot contains a lot of data, we should examine each dimension in turn. On the horizontal axis, dimension 1 contrasts *pretty* and *quite*. Each of them attracts

25. In CA, “inertia” is very similar to the “moment of inertia” in applied mathematics. It measures the total variance of the data table.

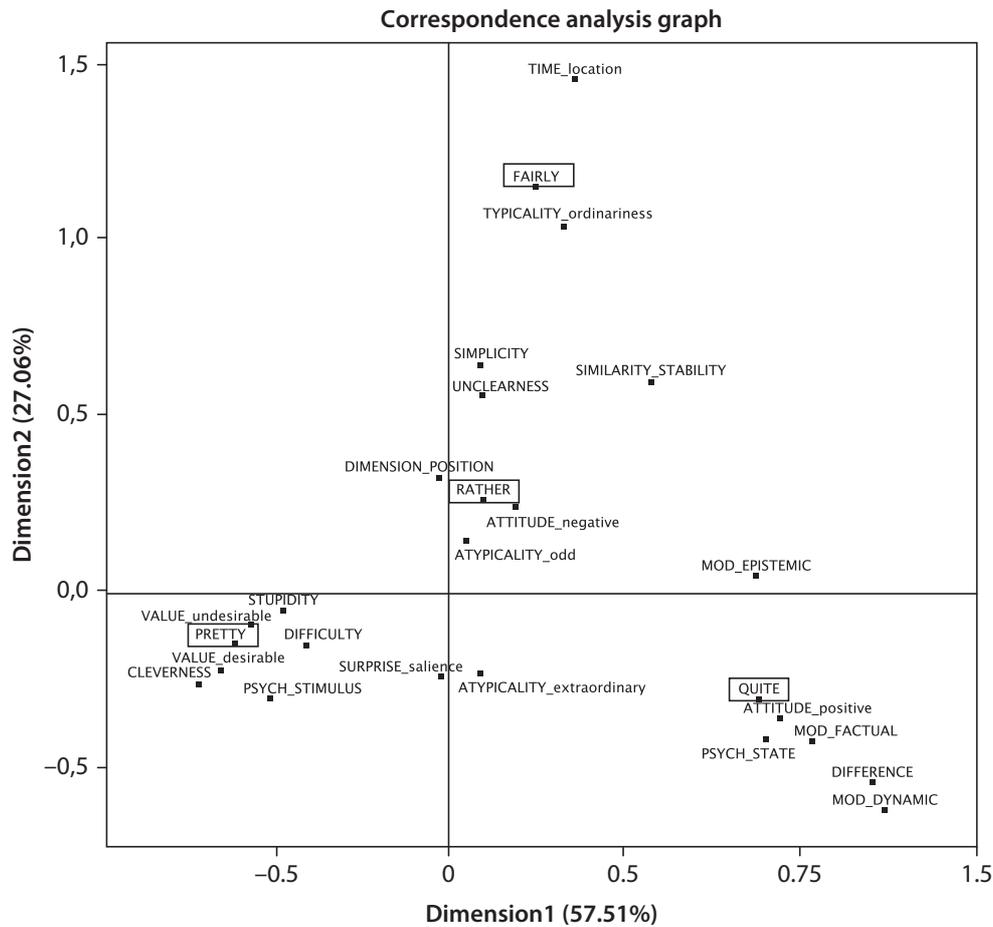


Figure 3. CA biplot of the <moderator + adjective> construction in COCA (with semantic annotation)

its own cloud of adjectives, and each cloud is clearly delimited. On the vertical axis, dimension 2 opposes *fairly* and *rather* (at the top of the cloud) to *pretty* and *quite* (at the bottom of the cloud). This goes against the implicit assumption that moderators split up between *rather* and *quite* on the one hand, and *pretty* and *fairly* on the other hand (see, for example, Downing & Locke 2006). The proximity between *fairly* and *rather* is evidenced by the continuum formed by their distinctive collexemes (both horizontally and vertically). Comparatively, only two adjectives (*surprising* and *difficult*) stand halfway between *rather* and *quite*. In all likelihood, the boundary between *fairly* and *rather* can be drawn above a cluster of adjectives with negative connotations (*obscure*, *crude*, *mundane*, *lengthy*, *simplicistic*), which are distinctive of *rather*. At this stage, it is still difficult to spot any division of labor among moderators because of the granularity of the plot. Figure 3, above, presents the graphic output of CA once all adjectives have been semantically annotated. The relative position of moderators in the cloud is very similar to the configuration displayed in Figure 2.

Annotating adjectives makes it considerably easier to identify the functional specificities of each moderator as well as the division of labor among them. The specificities of each moderator are listed below:

- rather*: dimension or position in space (ex. *long, high*), atypicality/oddity (ex. *odd, bizarre*), negative attitudes (ex. *ironic*), unclearness (ex. *vague, obscure*);
- quite*: epistemic, dynamic, and factual meanings (ex. *likely, able, true*), difference (ex. *different, separate*), psychological states (ex. *surprised, concerned, content*);
- fairly*: location in time (ex. *recent, new*), typicality (ex. *typical, common, standard*);
- pretty*: appreciative and unappreciative values (ex. *good, great* vs. *bad, awful*), cleverness and stupidity (ex. *smart* vs. *stupid, dumb*), difficulty (ex. *difficult, tough, hard*), psychological stimuli (ex. *scary, funny*).

The above list shows that moderators follow a division of labor in the intensification of some complementary meanings:

- *rather* modifies spatial location and atypicality whereas *fairly* modifies time location and typicality;
- *rather* modifies the expression of negative attitude whereas *quite* modifies the expression of positive attitude;
- *pretty* modifies the expression of difficulty whereas *fairly* and *rather* modify the expression of simplicity;
- *pretty* modifies the expression of psychological stimuli whereas *quite* modifies the expression of psychological states;
- *quite* modifies the expression of difference whereas *fairly* and *rather* modify the expression of similarity/stability.

Lastly, some meanings are not distinctive of any moderator in particular:

- modifying the degree of surprise/salience and atypicality/extraordinariness is common to *pretty, rather, and quite*;
- modifying the degree of simplicity and similarity/stability is common to both *fairly* and *rather*.

To summarize, we have three major configurations:

- first configuration: moderators operate within one conceptual content;
- second configuration: two complementary aspects of a conceptual domain are intensified by two distinct moderators;
- third configuration: one conceptual content can be intensified indiscriminately by different moderators.

5. Discussion and conclusion

In this paper, we have proposed and combined several statistical methods to provide a bidirectional semantic modeling of the <moderator + adjective> construction. We have made three points. Firstly, we have reasserted the need for better statistics in the collocation-based study of degree modifiers. Collostructional analysis is superior to most techniques based on raw counts and/or percentages because it filters away co-occurring pairs that are unrealistically too frequent or too rare, regardless of the size of the corpus. Secondly, we have shown that combining univariate and multivariate statistics can help map usage patterns and conceptual structure in a set of near-synonyms. The relationship between moderators and adjectives is indeed bidirectional, and it can be represented spatially. Thirdly, my results partly support Paradis (1997) regarding the cognitive synonymy of moderators, which are both similar and different. Moderators are similar because they have a functional basis in common, namely modifying the degree of a property denoted by an adjective. Moderators are also different because they do not modify the same classes of adjectives. In Cognitive Grammar terms, moderators do not always operate within the same conceptual domains. If they do, they follow a division of labor.

These findings are of great significance to the study of constructions since two items that co-occur significantly are likely to be entrenched as a constructional unit.²⁶ Figure 2 shows that some pairs are more entrenched than others. For example, *quite surprised* is more entrenched than *quite surprising*; *pretty crazy* is more entrenched than *pretty silly*; *rather vague* is more entrenched than *rather abstract*; and *fairly straightforward* is more entrenched than *fairly easy*. Once a pairing of lexemes is sufficiently entrenched, it is likely to acquire a meaning/function of its own. Figure 3 shows that the division of labor of moderators is not limited to the expression of intensification. It includes the expression of various meanings, such as the expression of modality, value judgments, dimension, position in time or in space, etc.

Cognitive Construction Grammar takes an inventory approach to the mental representation of grammar. Such an approach assumes that grammar predominantly stores language structure in a complex constructional network instead of building structure “on demand”. In a section on partial productivity, Goldberg postulates that the candidates for the verb slot in the ditransitive construction are stored in speakers’ memories as similarity clusters on the basis of their type frequencies (1995: 133–136). The higher the type frequency, the bigger the cluster, and the more productive the

26. However we should be wary of not establishing too strong a correspondence between high frequencies and entrenchment (Geeraerts 2000). Studies in cognitive semantics have shown that what determines the entrenchment of a linguistic unit is not so much its high frequency as its absolute frequency as its frequency of occurrence relative to the frequency of similar units in similar contexts (Geeraerts, Grondelaers & Bakema 1994). Also, some linguistic units are entrenched not because they occur frequently, but because they are salient (Schmid 2007, 2010).

verb class is. In an effort to map usage patterns, Goldberg provides a two-dimensional representation where verbs cluster spatially according to similarity (1995: 135). Accordingly, *give*, *pass*, *bequeath* or *grant* are good candidates for the verb slot in the ditransitive construction, whereas *envy* or *forgive* are poor (but by no means impossible) candidates. Goldberg's map is a theoretical abstraction because it is not based on actual corpus data or similarity metric, as opposed to Figure 2 and Figure 3 in Section 4. Although the latter bear on a different case study, they can be considered as corpus-driven and statistically grounded extensions of Goldberg's graphic intuition. Figure 2 is flexible enough to represent both entrenched collocations (e.g. *rather vague*, *quite different*, *fairly new*, *pretty good*) and collocations that are improbable, yet possible (e.g. *rather neat*, *quite cool*, *fairly stupid*, *pretty right*). It reflects the fact that speakers tend to use certain adjectives with certain degree-modifiers, but can also extend moderators idiosyncratically to other classes of adjectives. The existence of dense, neat clusters of adjectives around *pretty* and *quite* suggests that speakers are more conservative in their use of adjectives with these two moderators. Figure 3 confirms that adjectives cluster around moderators on the basis of semantic similarity. In sum, I have presented evidence that shows that types of the <moderator + adjective> construction form a network structured by similarity clusters.

Recent studies on near-synonymy in the Cognitive Linguistic framework have concluded that multifactorial techniques can help map usage patterns (Glynn 2010b) and spot "clusters in the mind" (Divjak & Gries 2008; Divjak 2010). Given how representative the corpus I have used is, the same kind of conclusions can be drawn regarding the <moderator + adjective> construction, pending experimental verification.

Perhaps in comparison to other less macroscopic approaches, the results presented in this paper seem conditional. Nevertheless, these results take usage-based representations of linguistic units seriously and are verifiable. One can test these findings with a different corpus and compare the results. Confirmatory statistics such as logistic or log-linear regression will have to corroborate the claim that these findings are not due to chance and provide a faithful representation of the reality of the data. For reasons of space, I have deliberately left aside three aspects of the <moderator + adjective> construction, some of which have received much attention in the past, such as syntactic idiosyncrasies (Allerton 1987; Gilbert 1989), grading force (Paradis 1997), and subjectivity (Nevalainen & Rissanen 2002; Athanasiadou 2007). However, I believe that the methodology presented in this paper can shed new light on each of these aspects. Regarding syntactic idiosyncrasies, distinctive collexemes can be used as input for correspondence analysis to obtain a clearer picture of alternations such as <*quite a* + adjective> vs. <*a quite* + adjective>, <*quite*_{MODERATOR} + scalar adjective> vs. <*quite*_{MAXIMIZER} + absolute adjective>, or <*a fairly* + adjective + noun> vs. <NP + *be* + *fairly* + adjective>. Regarding grading force, adjectives can be annotated according to gradability in correspondence analysis following the categories proposed in Paradis (1997: 49), namely non-gradable, scalar, extreme, limit. Regarding subjectivity, the methodology I have proposed can be applied to a diachronic corpus (see

also Hilpert 2006), so as to conduct a quantitative assessment of subjectification over the long-term history of *rather*, *quite*, *fairly*, and *pretty*.

I have made use of multifactorial methods to visualize data that is not properly speaking multifactorial. Instead of clustering lexical co-occurrence data alone, it will be in the interest of future research to expand what I have presented in Figure 3 and integrate more strata of richly annotated data within the same plot (e.g. information concerning grading force, boundedness, and semantic classes of adjectives). The resulting two-dimensional map will provide a finer-grained representation of the scale of synonymy of moderators. It will also help explain why collocations featuring the same adjective do not have the same connotation depending on which moderator is used.²⁷

To summarize, using multiple distinctive collexemes as input for correspondence analysis has three assets. First, it enables the linguist to ignore collocates that would be frequent or infrequent whatever the context (because they have a high overall frequency throughout the corpus) and focus on relevant pairs. Once distinctive collexemes have been identified, their raw frequencies can be used safely in a multi-way table to map correlations between moderators and adjectives by means of a multivariate statistical technique.

Second, one graphic output is enough to synthesize similarities and differences in a set of near-synonyms. Similarities between moderators (e.g. *rather* and *fairly*) are evidenced by their relative proximity on the map. Differences (e.g. *pretty* vs. *quite*) are made apparent by the relative distance between items. This is also true of adjectives, which tend to cluster according to meaning.

Third, visualizing distinctive collexemes in a correspondence analysis plot can do more than depict proximities and distances within separate paradigms. It is also a potentially accurate means of determining entrenchment continua along the two dimensions that structure the Euclidean space. Hopefully, the methodology I have proposed can be used to represent the complex inventory of constructions that shapes speakers' grammars.

References

- Allerton, D. J. (1987). English intensifiers and their idiosyncrasies. In R. Steele & T. Threadgold (Eds.), *Language topics: Essays in honour of Michael Halliday* (pp. 15–31). Amsterdam & Philadelphia: John Benjamins.

27. For example, even though *good* is a distinctive collexeme of *pretty*, raw frequencies show that it also co-occurs with *fairly* and *quite*. Presumably, it does not have the same meaning in each construction. It will be in the interest of future research to include context-based variables to clarify these variations in meaning.

- Altenberg, B. (1991). Amplifier collocations in spoken English. In S. Johansson & A. Stenström (Eds.), *English computer corpora: Selected papers and research guide* (pp. 127–149). Berlin & New York: Mouton de Gruyter.
- Athanasiadou, A. (2007). On the subjectivity of intensifiers. *Language Sciences*, 29, 554–565. DOI: 10.1016/j.langsci.2007.01.009
- Benzécri, J.-P. (1973). *L'analyse des données, 2. L'analyse des correspondances*. Paris: Dunod.
- Benzécri, J.-P. (1984). *Analyse des correspondances, exposé élémentaire* (2nd ed.). Paris: Dunod.
- Biber, D., Conrad, S., & Reppen, R. (1998). *Corpus linguistics – Investigating language structure and use*. Cambridge: Cambridge University Press. DOI: 10.1017/CBO9780511804489
- Bolinger, D. L. M. (1972). *Degree words*. The Hague: Mouton. DOI: 10.1515/9783110877786
- Church, K. W., & Hanks, P. (1990). Word association norms, mutual information, and lexicography. *Computational Linguistics*, 16, 22–29.
- Cruse, D. A. (1986). *Lexical semantics*. Cambridge: Cambridge University Press.
- Davies, M. (1990–present). *The Corpus of Contemporary American English (COCA): 410 + million words*. <http://corpus.byu.edu/coca>
- Divjak, D. (2006). Ways of intending: Delineating and structuring near-synonyms. In St. Th. Gries & A. Stefanowitsch (Eds.), *Corpora in Cognitive Linguistics: Corpus-based approaches to syntax and lexis* (pp. 19–56). Berlin & New York: Mouton de Gruyter.
- Divjak, D. (2010). *Structuring the lexicon: A clustered model for near-synonymy*. Berlin & New York: Mouton de Gruyter.
- Divjak, D., & Gries, St. Th. (2006). Ways of trying in Russian: Clustering behavioral profiles. *Corpus Linguistics and Linguistic Theory*, 2, 23–60. DOI: 10.1515/CLLT.2006.002
- Divjak, D., & Gries, St. Th. (2008). Clusters in the mind? Converging evidence from near synonymy in Russian. *The Mental Lexicon*, 3, 188–213. DOI: 10.1075/ml.3.2.03div
- Downing, A., & Locke, P. (2006). *English grammar: A university course* (2nd ed.). London: Routledge.
- Edmonds, P., & Hirst, G. (2002). Near-synonymy and lexical choice. *Computational Linguistics*, 28, 105–144. DOI: 10.1162/089120102760173625
- Everitt, B. S., Landau, S., Leese, M., & Stahl, D. (2011). *Cluster Analysis* (5th ed.). Oxford: Wiley-Blackwell. DOI: 10.1002/9780470977811
- Evert, S. (2004). The statistics of word cooccurrences: Word pairs and collocations. Unpublished doctoral dissertation. Institut für maschinelle Sprachverarbeitung. University of Stuttgart.
- Firth, J., R. (1957). A synopsis of linguistic theory, 1930–1955. In J. R. Firth (Ed.), *Studies in linguistic analysis*. Special volume of the Philological Society (pp. 1–32). Oxford: Blackwell.
- Geeraerts, D. (2000). Salience phenomena in the lexicon: A typology. In L. Albertazzi (Ed.), *Meaning and Cognition* (pp. 79–101). Amsterdam & Philadelphia: John Benjamins.
- Geeraerts, D., Grondelaers, S., & Bakema, P. (1994). *The structure of lexical variation: Meaning, naming, and context*. Berlin, New York: Mouton de Gruyter.
- Gilbert, E. (1989). Quite, rather. *Cahiers de recherche en grammaire anglaise*, 4, 4–61.
- Gilquin, G. (2007). The verb slot in causative constructions. Finding the best fit. *Constructions*, SV1-3/2006. www.elanguage.net/journals/index.php/constructions/article/view/18/23
- Glynn, D. (2010a). Corpus-driven Cognitive Semantics. An introduction to the field. In D. Glynn & K. Fischer (Eds.), *Corpus-driven Cognitive Semantics: Quantitative approaches* (pp. 1–42). Berlin & New York: Mouton de Gruyter. DOI: 10.1515/9783110226423.1
- Glynn, D. (2010b). Synonymy, lexical fields, and grammatical constructions. A study in usage-based Cognitive Semantics. In H. Schmid & S. Handl (Eds.), *Cognitive foundations of linguistic usage-patterns: Empirical studies* (pp. 89–118). Berlin & New York: Mouton de Gruyter.

- Goldberg, A. E. (1995). *Constructions: A construction grammar approach to argument structure*. Chicago: University of Chicago Press.
- Goldberg, A. E. (2003). Constructions: A new theoretical approach to language. *Trends in Cognitive Sciences*, 7, 219–224. DOI: 10.1016/S1364-6613(03)00080-9
- Goldberg, A. E. (2006). *Constructions at work: The nature of generalization in language*. Oxford: Oxford University Press.
- Goldberg, A. E. (2009). The nature of generalization in language. *Cognitive Linguistics*, 20, 201–224. DOI: 10.1515/COGL.2009.013
- Greenacre, M. J. (2007). *Correspondence analysis in practice* (2nd ed.). Boca Raton: Chapman & Hall/CRC. DOI: 10.1201/9781420011234
- Gries, St. Th. (2007). Coll.analysis 3.2. A program for R for Windows 2.x.
- Gries, St. Th. (2010). *Statistics for linguistics with R: A practical introduction*. Berlin & New York: Mouton de Gruyter.
- Gries, St. Th., & Stefanowitsch, A. (2004). Extending collocation analysis: A corpus-based perspective on ‘alternations’. *International Journal of Corpus Linguistics*, 9, 97–129. DOI: 10.1075/ijcl.9.1.06gri
- Gries, St. Th. & Stefanowitsch, A. (2010). Cluster analysis and the identification of collexeme classes. In J. Newman, & S. Rice (Eds.), *Empirical and experimental methods in cognitive/functional research* (pp. 73–90). Stanford, CA: CSLI.
- Gries, St. Th. & Stefanowitsch, A. (2006). *Corpora in Cognitive Linguistics: Corpus-based approaches to syntax and lexis*. Berlin & New York: Mouton de Gruyter. DOI: 10.1515/9783110197709
- Hilpert, M. (2006). Distinctive collexeme analysis and diachrony. *Corpus Linguistics and Linguistic Theory*, 2, 243–256. DOI: 10.1515/CLLT.2006.012
- Hirst, G. (1995). Near-synonymy and the structure of lexical knowledge. In J. Klavans (Ed.), *AAAI Symposium on Representation and Acquisition of Lexical Knowledge: Polysemy, Ambiguity, and Generativity* (pp. 51–56). Cambridge (Mass.): AAAI Press.
- Kay, P. (2013). The limits of (Construction) Grammar. In T. Hoffmann & G. Trousdale (Eds.), *The Oxford Handbook of Construction Grammar* (pp. 32–48). Oxford: Oxford University Press.
- Kennedy, G. (2003). Amplifier collocations in the British National Corpus: Implications for English language teaching. *TESOL Quarterly*, 37, 467–487. DOI: 10.2307/3588400
- Kilgarriff, A. (2001). Comparing corpora. *International Journal of Corpus Linguistics*, 6, 97–133.
- Kilgarriff, A. (2005). Language is never, ever, ever, random. *Corpus Linguistics and Linguistic Theory*, 1, 263–276.
- Langacker, R. W. (1987). *Foundations of Cognitive Grammar. Vol. 1. Theoretical prerequisites*. Stanford: Stanford University Press.
- Langacker, R. W. (1991). *Foundations of Cognitive Grammar. Vol. 2. Descriptive application*. Stanford: Stanford University Press.
- Langacker, R. W. (2008). *Cognitive Grammar: A basic introduction*. Oxford: Oxford University Press.
- Langacker, R. W. (2009). Cognitive (construction) grammar. *Cognitive Linguistics*, 20, 167–176.
- Leech, G. (1997). Introducing corpus annotation. In R. Garside, G. Leech, & T. McEnery (Eds.), *Corpus annotation: Linguistic information from computer text corpora* (pp. 1–18). London: Longman.

- Lorenz, G. (2002). Really worthwhile or not really significant? A corpus-based approach to the delexicalization and grammaticalization of intensifiers in Modern English. In I. Wischer, & G. Diewald (Eds.), *Speech, place, and action: Studies in deixis and related topics* (pp. 143–161). New York: Wiley.
- Manning, C. D., & Schütze, H. (1999). *Foundations of statistical natural language processing*. Cambridge, MA: MIT Press.
- Nevalainen, T., & Rissanen, M. (2002). Fairly pretty or pretty fair? On the development and grammaticalization of English downtoners. *Language Sciences*, 24, 359–380.
- Paradis, C. (1994). Compromisers – a notional paradigm. *Hermes*, 13, 157–167.
- Paradis, C. (1997). *Degree modifiers of adjectives in spoken British English*. Lund: Lund University Press.
- Paradis, C. (2000). It's well weird: Degree modifiers of adjectives revisited: The nineties. In J. M. Kirk (Ed.), *Corpora galore: Analyses and techniques in describing English* (pp. 147–160). Amsterdam & Atlanta: Rodopi.
- Paradis, C. (2008). Configurations, construals and change: Expressions of degree. *English Language and Linguistics*, 12, 317–343.
- Pedersen, T. (1996). Fishing for exactness. In *Proceedings of the South-Central SAS Users Group Conference* (pp. 188–200). Austin, TX.
- Quine, W. V. O. (1951). Main trends in recent philosophy: Two dogmas of empiricism. *The Philosophical Review*, 60, 20–43.
- Quirk, R., Greenbaum, S., Leech, G., & Svartvik, J. (1985). *A comprehensive grammar of the English language*. London & New York: Longman.
- R Development Core Team. (2011). *R: A language and environment for statistical computing*. Vienna: Austria: R Foundation for Statistical Computing.
- Schmid, H.-J. (2007). Entrenchment, salience, and basic levels. In D. Geeraerts & H. Cuyckens (Eds.), *The Oxford handbook of Cognitive Linguistics* (pp. 117–138). Oxford: Oxford University Press.
- Schmid, H.-J. (2010). Does frequency in text instantiate entrenchment in the cognitive system? In D. Glynn & K. Fischer (Eds.), *Quantitative methods in cognitive semantics: Corpus-driven approaches* (pp. 101–133). Berlin, New York: Mouton De Gruyter.
- Simon-Vandenberg, A.-M. (2008). Almost certainly and most definitely: Degree modifiers and epistemic stance. *Journal of Pragmatics*, 40, 1521–1542.
- Sinclair, J. (1991). *Corpus, concordance, collocation*. Oxford: Oxford University Press.
- Stefanowitsch, A., & Gries, St. Th. (2003). Collostructions: Investigating the interaction of words and constructions. *International Journal of Corpus Linguistics*, 8, 209–243.
- Stoffel, C. (1901). *Intensives and downtoners: A study in English adverbs*. Heidelberg: Carl Winter.
- Storjohann, P. (2009). Plesionymy: A case of synonymy or contrast? *Journal of Pragmatics*, 41, 2140–2158.
- Traugott, E. C. (2008). The semantic development of scalar focus modifiers. In A. van Kemenade & B. Los (Eds.), *The handbook of the history of English* (pp. 335–359). Oxford: Blackwell.
- Ward, J. H. (1963). Hierarchical grouping to optimize an objective function. *Journal of the American Statistical Association*, 58, 236–244.
- Wiechmann, D. (2008). On the computation of collostruction strength: Testing measures of association as expressions of lexical bias. *Corpus Linguistics and Linguistic Theory*, 4, 253–290.