



HAL
open science

Forms and meanings of intensification: a multifactorial comparison of quite and rather

Guillaume Desagulier

► **To cite this version:**

Guillaume Desagulier. Forms and meanings of intensification: a multifactorial comparison of quite and rather. *Anglophonia / Caliban - French Journal of English Linguistics*, 2015, Degré et intensification, 20, <10.4000/anglophonia.558>. <halshs-01183658>

HAL Id: halshs-01183658

<https://shs.hal.science/halshs-01183658v1>

Submitted on 12 Aug 2015

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.



Distributed under a Creative Commons CC BY-NC-SA 4.0 - Attribution - Non-commercial use - ShareAlike - International License

Forms and meanings of intensification: a multifactorial comparison of *quite* and *rather*

Guillaume Desagulier

Université Paris 8, Université Paris Ouest / Nanterre La Défense, UMR 7114 MoDyCo
gdesagulier@univ-paris8.fr

Résumé

Pour cerner les relations entre quasi-synonymes dans une approche fondée sur l'usage, j'exploite des fréquences de collocation à l'aide de méthodes multifactorielles exploratoires. Mon étude se limite à quite et rather dans les contextes où ils intensifient des adjectifs dans le British National Corpus. J'ai recours à l'analyse factorielle des correspondances puis à l'analyse des correspondances multiples pour visualiser et interpréter des distances entre (a) les deux intensifieurs, (b) les adjectifs intensifiés ainsi que les classes sémantiques respectives auxquelles ils appartiennent et (c) les propriétés syntaxiques concernant l'agencement entre les intensifieurs et les adjectifs. Les résultats montrent que quite et rather appartiennent à un réseau cohérent de constructions. Dans ce réseau, les deux adverbes se répartissent les tâches dans l'intensification de sens adjectivaux. Le premier des deux résultats principaux de cette étude est le suivant : lorsqu'un adjectif a deux connotations, l'une positive et l'autre négative, il sera intensifié de préférence par rather lorsque sa connotation est négative. Le deuxième est que dans le strict cadre de l'alternance pre-déterminant vs. pré-adjectival, quite apparaît de préférence avant le déterminant et rather avant l'adjectif.

Abstract

To capture usage-based relations between near-synonyms, I cluster collocation data using exploratory multifactorial methods. My investigation is restricted to quite and rather in the contexts where they intensify adjectives in the British National Corpus. I use correspondence analysis and multiple correspondence analysis to visualize and interpret distances between (a) the two intensifiers, (b) the adjectives they modify and the respective semantic classes they belong to, and (c) syntactic information regarding how intensifiers and adjectives pattern together. Results show that quite and rather constructions form a consistent network. The first key finding is that they typically follow a division of labor in the intensification of adjectival meanings. When a positive and a negative connotation are available for a given adjective, rather tends to intensify the negatively connoted adjective. The second key finding is the following: in the strict frame of the pre-determiner vs. pre-adjectival alternation, quite displays a preference for the pre-determiner position, and rather for the pre-adjectival position.

Mots-clés : collocations, analyse collocationnelle, analyse factorielle des correspondances, analyse des correspondances multiples, intensifieurs

Keywords: collocations, collocational analysis, correspondence analysis, multiple correspondence analysis, intensifiers

1. Introduction

Usage-based models of language posit that grammar is acquired, represented mentally, and accessed in a ‘bottom-up’ fashion, through exposure to usage events (Bybee, 2006; Langacker, 1988; 1999; 2000). Because usage events are intrinsically context dependent, grammar is a structured inventory of symbolic units whose architecture is shaped by the accumulation of linguistic experience. Given this focus on context and this holistic approach to linguistic experience, a fast-growing community of cognitive linguists have begun to realize the assumptions of their own theoretical framework were essentially empirical and had to be tested (Gibbs, 2007; Gries, Hampe, & Schönefeld, 2005; Tummers, Heylen, & Geeraerts, 2005). As a result, two lines of research emerged at the crossroads of lexical semantics and construction grammar, the first one focusing on how sociolinguistic factors determine the choice of synonyms, and the second one relying on multifactorial methods to differentiate between near-synonyms (for an overview, see Geeraerts, 2010: 263-264). This paper belongs to the second research area.

I propose quantitative techniques to capture subtle usage-based relations between two near-synonyms in British English: *quite* and *rather*. Although *quite* and *rather* can modify other adverbs (*quite frankly*; *rather desperately*), noun phrases (*quite a sight*; *rather a shock*), or even verb phrases (*I quite understand*; *I rather enjoyed it*), my investigation is restricted to the far more frequent contexts where these intensifiers are used as degree modifiers of adjectives:

- (1) *We are **quite different** people now, we need different things.* (BNC-CEX)i
- (2) *Yes, and there’s a **rather odd** smell in this room...* (BNC-J17)

When *quite* and *rather* modify attributive adjectives, they can occur in pre-determiner position, a behavior that other intensifiers do not show:

- (3) a. *That has proved to be **a quite difficult** question to answer.*
b. *That has proved to be **quite a difficult** question to answer.*
- (4) a. *That **is a rather difficult** question to answer.*
b. *That **is rather a difficult** question to answer.*
- (5) a. *I know it’s **a fairly difficult** question.*
b. *??I know it’s **fairly a difficult** question.*

Allerton (1987: 25) believes that the choice of pre-determiner position over the default pre-adjectival position is more than a matter of style or formality. In this paper, my first goal is to use quantitative corpus-based techniques to decide if the difference between the pre-adjectival and the pre-determiner patterns of *quite* and *rather* translates into a difference in meaning.

Allerton observes further that some restrictions apply depending on whether the adjective that *quite* modifies is scalar or absolute:

- (6) a. *I mean this is **quite a good idea** / **?a quite good idea** actually.*
b. *This is **?quite an excellent idea** / **a quite excellent idea**.*

More precisely, pre-determiner uses of *quite* and *rather* seem to inherit the emphatic and exclamatory meaning of <*quite/rather* + NP> (e.g. *quite a pen!*, *rather a fool!*).

Typically, when the intensifier precedes the determiner, it has a modal meaning and it functions as a sentence-modifier, not a word-modifier (Allerton 1987, Bolinger 1972, Stoffel 1901). Regarding *quite*, however, Bolinger (1972: 137) observes the following:

(...) somehow along the way, the indefinite article has ceased to separate the two functions consistently, with the result that *quite a* and *a quite*, for example, form an alternating pattern with so slight a difference in meaning that outside factors may decide the choice.

In the light of subsequent research on synonymy, Bolinger's observation cannot be taken for granted. Bolinger himself claims elsewhere that "if two ways of saying something differ in their words or their arrangement they will also differ in meaning" (1977: 1). In the same vein, Clark's Principle of Contrast stipulates that "every two forms contrast in meaning" (Clark 1987: 2, 1992: 172), and Goldberg's Principle of No Synonymy states that "[i]f two constructions are syntactically distinct, they must be semantically or pragmatically distinct" (Goldberg 1995: 67).

Admittedly, the semantic difference between the pre-adjectival and pre-determiner structures in (3) is slim. However, if we examine usage closely, pre-adjectival and pre-determiner uses do not overlap neatly. Indeed, a cursory search in the British National Corpus (World Edition) reveals that *quite* tends to intensify adjectives that denote importance in size (*big*, *considerable*, *sizeable*, *substantial*) when it precedes the determiner. There is no such preference when *quite* follows the determiner and precedes the adjective. Semantic differences between alternating patterns may be small, but not necessarily insignificant in terms of conventionally sanctioned usage. Therefore, my second, more general goal is to inspect the near synonymy of *quite* and *rather* with respect to form and meaning.

Most major studies on intensifiers have focused on the best way to infer gradability from collocational preferences (Altenberg 1991, Kennedy 2003, Lorenz 1999, 2002, Simon-Vandenberg 2008). One reason behind this methodological preference is the following: because the vast majority of degree modifiers of adjectives function as 'word modifiers' (Stoffel: 1901), they lend themselves to straightforward quantitative corpus analysis. However, the methodology used in those studies can be deemed faulty on (at least) two counts. Firstly, syntactic variables are seldom incorporated into the abovementioned collocation-based descriptions of adjective intensification (Lorenz 2002, Kennedy 2003). When they are (e.g. Lorenz [1999], Simon-Vandenberg [2008]), these variables are generally listed irrespective of their potential interaction with semantic features. Secondly, whether based on raw counts, coarse-grained relative frequencies such as percentages and counts per *n*-thousand words (e.g. Altenberg [1991] and Paradis [1997]), or far more reliable statistical association measures, collocation analysis cannot detect relevant patterns of interaction among multiple factors because it is meant for the exploration of one variable at a time. While the approach presented in this article shares substantial common ground with the above studies, it departs from them methodologically.

I will therefore use multifactorial techniques to examine the co-occurrence of intensifiers and adjectives in the context of the syntactic patterns where they occur. These patterns are the pre-determiner *vs.* pre-adjectival positions of *quite* and *rather* and the predicative *vs.* attributive positions of the adjectives that these adverbs intensify. Results will show that the focal adjustment on the meaning of adjectives does not depend exclusively on the choice of an adverb but also on the syntax of both adverbs and adjectives.

This paper is organized as follows. Section 2 reviews previous approaches to English intensifiers and highlights some issues in the collocation-based methodologies that these works implement. Section 3 addresses the limitations of previous collocation analyses, advocates the need to include collocation data in multidimensional tables, and proposes a new methodology to handle such tables. Sections 4 and 5 present and discuss the results. Section 6 concludes on methodological issues and implications.

2. Issues with previous research

From a Cognitive Grammar perspective (Langacker 1987, 1991, 2008), 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 construalⁱⁱ. If *quite* and *rather* are indeed near synonyms, we should expect them to share the same conceptual content but to differ in terms of construal. The conceptual content that *quite* and *rather* share is essentially functional: both intensifiers index the properties denoted by the adjectives they modify on a scale ('upwards' or 'downwards'). The fact that both *quite* and *rather* alternate between pre-determiner and pre-adjectival positions is a sign that their functions overlap.

Early signs of construal differences between *quite* and *rather* can be traced back to their respective etymologies. *Quite* is borrowed from Anglo-Norman *quit/quite* ('without opposition'), which can be traced back to Latin *quietus* ('at rest'). Its original meaning is the absence of obligation, the discharge of one's duty. We may assume that 'freedom from coercion' grammaticalized into 'absence of grading restriction' by the time *quite* was first attested in Middle English, because it was then used as a boosting intensifier. 'Absence of grading restriction' further shifted to 'presence of a property to some degree' in the early 19th century, when *quite* started to be used as a moderating intensifier:

- (7) *Had Lawyer L--n staid at home, His honour might have pass'd, with some, For quite a decent country Squire.* (1806, T.G. Fessenden, *Democracy Unveiled* II. Vi. 204) (*OED online*).

The diachrony of *quite* offers valuable clues as to how the adverb intensifies adjectives. If we posit that an adjective denotes a property *P*, *quite* signals the following interpretation path: $\neg P$ is considered and rejected in favor of *P*, which is then indexed on an intensity scale. This is consistent with Bolinger (1972: 105), quoted by Gilbert (1989: 7-8), who observes that *quite cold* and *very cold* do not have the same connotations:

- (8) *It is quite cold this morning.*
 (9) *It is very cold this morning.*

He writes: "the first connotes an unexpectedness that the second lacks; it would be used of a cold day in summer, for example."

Originally, *rather* combines the adverb *rathe* 'quickly, rapidly' and the comparative suffix *-er*. It develops from later Middle English onwards, via the decline of *rathe*. The *OED online* states that '[t]he modifying function (...) probably arose from the reanalysis of the adverb's syntactic role in sentences where contrast is anaphorically implied (...)'. The idea of contrast is present in (8):

- (10) *He was very negligent himselfe & of a Philosophic temper & was indeede **rather negligent** of his person.* (a1684, *Diary* anno 1675 (1955), J. Evelyn, IV. 60) (*OED online*)

P ('negligent') and $\neg P$ ('not negligent') are compared and P is considered prevalent to some degree. The comparative nature of *rather* suggests that the intensifier selects P more than $\neg P$. This is consonant with Allerton's description of *rather* as a sentence adverbial in prosodically isolated positions:

- (11) a. *She's an **attractive** woman(,) rather.*
b. *She is – rather – an **attractive** woman.* (Allerton 1987: 27)

Here, *rather* signals that the use of the tonic adjective *attractive* is preferred to the use of another adjective.

There is more to the meaning of an intensifying construction than the meaning of the intensifier. Paradis (1997: 62) argues that the semantic relation between modifiers and adjectives is bidirectional: "(...) the adjective selects a degree modifier which in turn constrains the conceptualization of the gradability of the adjective definitively (...)". We can interpret the bidirectionality hypothesis in two ways. One way is to assume that intensifying constructions are assembled in a linear way, the adjective being selected first and then modified in an online, *post hoc* fashion by the adverb. This could explain why some adjectives, because of their semantics, do not combine easily with all intensifiers, e.g. [?]*quite curious*, [?]*rather perfect*ⁱⁱⁱ. However, from a usage-based perspective, adjective and intensifier are more likely to be selected simultaneously. No matter which interpretation of the bidirectionality hypothesis we choose, it does not take into consideration the influence of the syntactic context where adjectives and intensifiers combine.

Altenberg (1991) is well aware of the syntactic restrictions governing the use of boosters and maximizers (1991: 128-129) but he does not discuss them in depth. Instead, he focuses on the collocational preferences of amplifiers regardless of the phrasal contexts (1991: 130). Both Lorenz (2002) and Kennedy (2003) adopt lexeme-centered approaches. In contrast, Lorenz (1999: 201-211) devotes a whole section to adjective intensification, information structure, and phrasal environment in a comparison of English learners' tendencies and native speakers' preferences in four corpora. He observes that adjective intensification occurs naturally with rhematic adjectives in predicative position to the detriment of attributive occurrences. However, Lorenz does not correlate the predicative and attributive positions with the semantic classes that the different adjectives fall into. Simon-Vandenberg (2008) proceeds likewise. While her comparative study of *certainly* and *definitely* encompasses several kinds of distinctive variables (register, genre, dialogic context, and information structure), she does not examine their correlation. In sum, previous quantitative studies on degree modifiers of adjectives do not systematically incorporate syntactic variables in their collocational description of intensification. When they do, these variables are generally described irrespective of their potential interaction with semantic features.

Another issue is the choice of association measures to assess collocational preferences, namely raw frequencies (Lorenz 2002), basic standardized figures (Altenberg 1991, Lorenz 1999, Paradis 1997, Simon-Vandenberg 2008), and

		syntax of the adjective	
		attributive	predicative
syntax of the intensifier	<i>quite</i>	pre-adjectival	<quite ADJ NP> <DET quite ADJ NP>
		pre-determiner	<quite DET ADJ NP>
	<i>rather</i>	pre-adjectival	<rather ADJ NP> <DET rather ADJ NP>
		pre-determiner	<rather DET ADJ NP>

pointwise mutual information (Kennedy 2003). Such measures do not filter away lexemes that are unrealistically too frequent or too rare regardless of the words they collocate with. Based on raw frequencies in the BNC (World Edition), the top collocate of *quite* and *rather* in pre-adjectival position is *different* (1 256 and 477 tokens

Table 1. How *quite* and *rather* pattern with adjectives

respectively). Considering these collocation frequencies alone, one could easily jump to the conclusion that the degree of mutual association between *different* and both intensifiers is high. That would be overlooking the fact that *different* ranks among the top five adjectives in the whole corpus. If we compare these collocation frequencies to the overall frequencies of *different* (47 209 tokens), *quite* (16 087), and *rather* (5 521), we realize that *different* is distinctive of *quite* more than it is distinctive of *rather*.

3. Methods

Assuming that both intensifiers and adjectives contribute to the meaning of the intensifying construction, a good place to start is to inspect co-occurrences of *quite/rather* and the adjectives that they modify in a large corpus and determine preferred collocations with a reliable association measure. However, because speakers do not make lexical choices regardless of morphosyntactic and semantic/pragmatic considerations, it seems implausible to assess near synonymy accurately on the basis of lexical collocation alone, as argued extensively by Divjak (2006, 2010). Accordingly, I investigate the co-occurrence of intensifiers and adjectives in the context of the syntactic patterns where they occur. *Quite* and *rather* manifest themselves across eight patterns, which are summarized in Table 1. I contrast these alternations in their respective adjectival preferences using a method from collostructional analysis (Hilpert 2008, Stefanowitsch 2013), namely multiple distinctive collexeme analysis, henceforth MDCA (Gilquin 2007, Hilpert 2006, Desagulier 2014). Collostructional analysis is a family of methods aimed at measuring the degree of attraction or repulsion that words exhibit to constructions. It consists of collexeme analysis (Stefanowitsch & Gries 2003), distinctive collexeme analysis (Gries & Stefanowitsch 2004b), and co-varying collexeme analysis (Gries & Stefanowitsch 2004a, Stefanowitsch & Gries 2005, Desagulier 2015).

MDCA is a subtype of distinctive collexeme analysis, which is well suited for the purpose of finding subtle differences between pairs of near-equivalent constructions (e.g. the ditransitive *vs.* prepositional dative alternation). MDCA extends distinctive collexeme analysis to more than two constructions. It determines the probability of each lexeme’s observed frequency given its expected frequency in each construction. This probability is log-transformed and the resulting value captures distinctiveness (*pbin*). The co-occurrence between a lexeme and a construction is statistically significant if *pbin* is higher than 1.3 ($p < 0.05$). MDCA outputs a table that contains as many lines as

there are distinctive collexemes. In addition to observed frequency, expected frequency, and *pbin*, the columns provide two values to help the linguist identify meaningful patterns of attraction: *SumAbsDev* and *LargestDev*. *SumAbsDev* gives the sum of all absolute *pbin* values for a particular collexeme – the higher the figure, the more the adjective deviates from its expected distribution. *LargestDev* indicates which construction each lexeme is a distinctive collexeme of.

Thanks to MDCA, one can determine the probability of each adjective's observed frequency given its expected frequency in each pattern. Only those adjectives whose absolute distinctiveness values are significant are kept, along with their respective observed frequencies in each of the eight *quite/rather* patterns.

The obtained contingency table serves as the basis for the constitution of a multidimensional dataset of non-negative ratio-scaled data. This dataset contains supplementary frequency data regarding the syntax and the semantics of the *quite/rather* constructions. Supplementary columns provide the frequency of co-occurrence between each adjective type and the following variables:

- the syntactic position of the intensifier: pre-adjectival vs. pre-determiner;
- the syntactic position/function of the adjective: attributive vs. predicative;
- the sum total for each intensifier: *quite* vs. *rather*;
- the text mode: spoken vs. written.

Supplementary rows group the frequencies of adjectives according to 59 semantic classes listed in the appendix. The semantic classes are inspired from Dixon & Aikhenvald (2004). Classes that match the data better are added in an *ad hoc* fashion.

The dataset is then submitted to correspondence analysis (henceforth CA), an exploratory statistical technique that takes the frequencies of multi-way tables as input, then summarizes and visualizes distances between the variables (Benzécri 1973, 1984, Greenacre 2007, Glynn 2014). More precisely, CA uses the frequencies of the dataset to:

- compute a matrix based on χ^2 distance and determine the probability of global association between rows and columns;
- transpose the multidimensional distances to a two-dimensional plane that maps the correlations between the variables; each row and each column is thus represented as a point in the Euclidean space.

The graph is used to visualize relative distances between (a) intensifiers, (b) adjectives and their respective semantic classes, (c) syntactic information regarding how intensifiers and adjectives pattern together.

The data are taken from the British National Corpus (World Edition), which consists of 100 467 090 words of spoken and written British English divided among over 4 054 texts. The spoken component contains approximately 10 million words, and consists of transcripts of spontaneous conversation and context-governed recording samples. The written component contains approximately 90 million words, and consists of samples of many kinds of text material (newspapers, fiction books, unpublished memoranda, etc.) The whole corpus spreads across the period from 1960 to the early 1990s (Burnard 2000).

As opposed to most spoken corpora, the BNC is a relatively large corpus. Its size and its sampling scheme increase the reliability and validity of our observations. As opposed to other editions, the world edition of the BNC is freely and publicly available. It is annotated for parts of speech and extractions are done through an online query interface (Davies 2004). One major disadvantage is that the interface offers no way of

exporting query results for further edition with a text editor or spreadsheet software. Exporting para- or extralinguistic information can only be done manually (e.g. *via* copy and paste), which is not a viable option when the construction under investigation has a high token frequency.

4. Results

For each of the eight alternating patterns listed in Table **Erreur ! Signet non défini.**, all adjectival collocates were extracted from the BNC, amounting to 3,374 adjective types distributed across 28,790 construction tokens. To test whether the difference between the observed frequency and the expected frequency of each adjective in each construction was statistically significant, I performed an MDCA using *Coll.analysis 3.2* (Gries 2007).

Usually, lexical semanticists gather enough information from the top 10 or 20 collexemes and can safely ignore the rest of the output table (e.g. Hilpert 2008). That would be irrelevant here because the ten adjectives with the highest *pbin* and *SumAbsDev* values are distinctive of only one construction (Table 2). Therefore, the only information we can glean concerns the preference of <quite ADJ> for adjectives that denote modality in the broad sense – factual (*clear, right, true*), epistemic (*possible, likely, sure*) and dynamic (*prepared, capable, easy*) – and a positive mental disposition (*happy*). A more comprehensive way of inspecting the results is to focus on distinctiveness values for each of the 8 constructions (Table 3), but then the mass of information makes it hard to generalize over the table.

Coll_Word	obs freq	exp freq	pbin	SumAbsDev	LargestDev
<i>sure</i>	716	402.84	173.42	313.52	0_quite_adj_0
<i>clear</i>	570	325.52	124.56	227.21	0_quite_adj_0
<i>happy</i>	543	313.20	111.76	208.21	0_quite_adj_0
<i>right</i>	407	242.04	70.89	143.63	0_quite_adj_0
<i>possible</i>	235	132.23	57.11	103.59	0_quite_adj_0
<i>likely</i>	178	100.85	41.28	74.80	0_quite_adj_0
<i>easy</i>	193	117.66	29.93	56.44	0_quite_adj_0
<i>prepared</i>	111	62.75	26.23	48.25	0_quite_adj_0
<i>capable</i>	120	68.91	25.77	47.91	0_quite_adj_0
<i>true</i>	107	61.07	23.85	43.90	0_quite_adj_0

Table 2. *Adjectives with the highest pbin and SumAbsDev values.*

Coll_Word	obs freq	exp freq	pbin	SumAbsDev	LargestDev
<i>sure</i>	716	402.84	173.42	313.52	0_quite_adj_0
<i>long</i>	227	24.35	168.36	289.18	quite_det_adj_np
<i>clear</i>	570	325.52	124.56	227.21	0_quite_adj_0
<i>happy</i>	543	313.2	111.76	208.21	0_quite_adj_0
<i>right</i>	407	242.04	70.89	143.63	0_quite_adj_0
<i>possible</i>	235	132.23	57.11	103.59	0_quite_adj_0
<i>different</i>	290	100.84	56.94	154.42	0_quite_adj_np
<i>good</i>	205	61.88	51.45	97	quite_det_adj_np
<i>big</i>	77	12.55	39.93	62.73	quite_det_adj_np
<i>large</i>	99	22.31	36.6	91	quite_det_adj_np
<i>distinct</i>	41	8.08	17.53	39.86	0_quite_adj_np
<i>separate</i>	33	7.09	13.08	36.09	0_quite_adj_np
<i>other</i>	15	1.31	13.03	26.94	0_quite_adj_np
<i>tired</i>	28	6.5	12.94	21.42	0_rather_adj_0
<i>extraordinary</i>	19	2.03	12.87	24.83	det_quite_adj_np
<i>surprised</i>	45	14.71	12.56	24.22	0_rather_adj_0
<i>substantial</i>	23	3.89	12.09	32.65	quite_det_adj_np
<i>unusual</i>	29	6.77	11.18	28.18	det_rather_adj_np
<i>slow</i>	29	7.74	11.14	24.35	0_rather_adj_0
<i>specific</i>	21	3.82	10.07	25.07	0_quite_adj_np
<i>remarkable</i>	18	2.66	9.73	16.96	det_quite_adj_np
<i>limited</i>	32	10.68	8.85	24.53	0_rather_adj_0
<i>gloomy</i>	11	1.45	7.76	18.59	det_rather_adj_np
<i>negative</i>	10	1.37	6.88	15.49	det_rather_adj_np
<i>eccentric</i>	10	1.37	6.88	14.23	det_rather_adj_np
<i>new</i>	13	2.16	6.61	19.27	det_quite_adj_np
<i>similar</i>	40	19.05	5.69	22.14	0_rather_adj_0
<i>grim</i>	8	1.13	5.47	13.24	det_rather_adj_np
<i>loud</i>	5	0.57	3.77	7.66	0_rather_adj_np
<i>inherent</i>	2	0.03	3.62	4.72	rather_det_adj_np
<i>recent</i>	5	0.63	3.46	10.58	det_quite_adj_np
<i>conservative</i>	4	0.41	3.32	7.91	0_rather_adj_np
<i>formal</i>	5	0.74	3.2	9.49	0_rather_adj_np
<i>fundamental</i>	4	0.48	2.94	10.54	det_quite_adj_np
<i>stronger</i>	3	0.25	2.9	6.99	0_rather_adj_np
<i>whole</i>	2	0.06	2.85	6.14	rather_det_adj_np
<i>low</i>	12	4.38	2.83	10.72	0_rather_adj_np
<i>essential</i>	2	0.08	2.64	4.88	rather_det_adj_np
<i>old</i>	5	1.39	1.88	6.6	rather_det_adj_np
...

Table 3. *Output of MDCA (sampled and sorted according to pbin value).*

If we focus on adverbs and disregard their syntax temporarily, some broad semantic tendencies emerge:

- *quite* constructions co-occur significantly with adjectives that denote modality (*sure, right, possible*), dimension or position in space (*big, large, substantial*), proximity in time (*new, recent*), singularity (*distinct*), and difference (*different, other*);
- *rather* constructions co-occur significantly with adjectives that denote oddities (*unusual, eccentric*) and retrograde leanings (*conservative, formal, old*);
- *quite* and *rather* contrast significantly in the expression of value judgments; *quite* attracts adjectives with a positive connotation (*happy, good, extraordinary, remarkable*) whereas *rather* attracts adjectives with a

negative connotation (*tired, slow, limited, gloomy, negative, eccentric, grim, loud, conservative, low*).

The last point is particularly striking. By itself, *surprised* does not have a negative connotation. Once modified by *rather* we may assume that the cause for surprise is unpleasant:

- (12) *My earliest memory is of standing in the nursery of Byron House in Highgate and crying my head off. (...) I think my parents were **rather surprised** at my reaction, because I was their first child and they had been following child development textbooks that said that children ought to start making social relationships at two.* (BNC-FYX)

Presumably, the same kind of negative bias is at work with adjectives such *formal, loud, or low^{iv}*:

- (13) *Joseph seemed to be curiously affected by the game. I was convinced it was not a romantic passion. I had **rather formal** notions of desire. Boys fancied you when your make-up was just right and your freckles didn't show and your hair was tidy.* (BNC-FU7)
- (14) *He was not short of that peculiarly British brand of slightly disdainful, **rather loud** self-confidence.* (BNC-CS4)
- (15) *Even the unemployed place government policy **rather low** in a list of factors responsible for high unemployment levels.* (BNC-B7G)

On the one hand, MDCA reveals coarse-grained semantic tendencies that would be beyond the reach of contingency tables based on raw frequencies. On the other hand, the above interpretation is partial because it is based on a small number of collexemes and it disregards how their meanings correlate with the syntax of the intensifiers. Insofar as MDCA is not designed for the cross-comparison of variables, the structural limitations of Table 3 render further exploration uneasy. To assess the constructional near synonymy of *quite* and *rather* more deeply and more systematically, we need to (a) inspect more construction tokens, (b) describe the data with syntactic and semantic variables that cover several levels of granularity, and (c) determine those variables beforehand (i.e. not heuristically). If we do so, we face another problem: as we include more data and more variables, we need to build and inspect several tables at once. In such conditions, making cross-comparisons is tedious and unproductive. Therefore, rather than inspect and compare many collocation-based frequency tables manually, I use correspondence analysis, a method that is well suited for the purpose of handling extensive datasets without compromising the complexity of the object of study.

From Table 3, we can build a contingency table that confronts two variables – the eight *quite/rather* constructions from Table 1 and their distinctive adjectives – taking advantage of the fact that the table offers a way to filter away those adjectives that are below a certain distinctiveness threshold (i.e. $p_{bin} > 1.3$, $p > 0.05$). However, studying the relationship between these two variables only is not particularly helpful for two reasons. First, my main objective is not just to see which adjectives pattern with which constructions. Rather, it is to determine to what extent the constructional profiles of *quite* and *rather* differ and to what extent they are similar. Second, the level of

granularity overviewed in Table 3 is too high. If we keep only those adjectives whose absolute distinctiveness value is higher than 5 we are potentially looking at 4 344 construction tokens (543 adjectives \times 8 constructions). This means that we should sort the data using categories that are broad enough to capture the general profiles of *quite* and *rather* and reflect the semantic and syntactic properties of each pattern.

To determine the semantic profile of a construction, I examine its distinctive adjectives, grouped by semantic classes. To determine the syntactic profile of a construction, I look at whether the intensifier occurs in pre-adjectival or pre-determiner position and whether the adjective is predicative or attributive. Given that the profile of a construction may be sensitive to text mode, we should also inspect variation according to the written vs. spoken distinction.

The above remarks translate into a dataset, sampled in Table 4. It contains:

- the raw frequencies of the 543 most distinctive adjectival collexemes ($pbin > 5$) of all 8 *quite/rather* constructions;
- 8 supplementary columns that correspond to the following variables: the syntactic position of the intensifier, the syntactic position/function of the adjective, the text mode, the sum total for each intensifier;
- 59 supplementary rows that group adjective frequencies according to semantic classes (positive value, negative value, atypicality, difference, etc.)

CA uses these frequencies to display the rows and columns as points on a two-dimensional map. Interpreting the geometric positions of the points is a way of interpreting the similarities and differences between rows (i.e. adjectives), the similarities and differences between columns (i.e. *quite* and *rather* constructions), and the association between rows and columns. CA is performed in R (R Core Team 2015) with the CA function of the `FactoMineR` package (Husson et al. 2009).

	X_QUITE_ADI_X	X_QUITE_ADI_NP	X_RATHER_ADI_X	X_RATHER_ADI_NP	DET_QUITE_ADI_NP	DET_RATHER_ADI_NP	QUITE_DET_ADI_NP	RATHER_DET_ADI_NP	PRE_DETERMINER_POSITION	PRE_ADJECTIVAL_POSITION	WITH_ATTRIBUTIVE_ADI	WITH_PREDICATIVE_ADI	SPOKEN	WRITTEN	QUITE_CONSTRUCTIONS	RATHER_CONSTRUCTIONS
<i>different</i>	770	290	213	97	174	160	136	8	144	1704	865	983	116	1732	1370	478
<i>difficult</i>	197	8	92	4	1	13	14	4	18	315	44	289	109	224	220	113
<i>nice</i>	301	13	68	4	3	10	72	7	79	399	109	369	356	122	389	89
<i>good</i>	509	70	64	11	6	17	205	25	230	677	334	573	550	357	790	117
<i>surprised</i>	47	0	45	3	0	0	0	0	0	95	3	92	34	61	47	48
<i>small</i>	137	46	43	11	6	17	33	7	40	260	120	180	39	261	222	78
<i>similar</i>	47	4	40	15	0	17	0	0	0	123	36	87	10	113	51	72
<i>strange</i>	25	0	38	9	1	20	1	2	3	93	33	63	23	73	27	69
<i>vague</i>	4	2	35	4	0	12	0	2	2	57	20	39	5	54	6	53
...
TOTAL (active rows)	12551	1313	2953	774	584	1509	1639	285	1924	19684	6104	15504	5229	16379	16087	5521
DIFFERENCE_contrast	1022	367	218	99	198	162	151	8	159	2066	985	1240	128	2097	1738	487
DIFFICULTY_complexity	364	54	160	15	18	55	46	9	55	666	197	524	215	506	482	239
DIMENSION_POSITION	966	291	343	126	55	192	642	79	721	1973	1385	1309	692	2002	1954	740
DISCOMFORT	16	1	31	11	0	34	5	0	5	93	51	47	17	81	22	76
DULLNESS	59	3	115	39	1	64	2	4	6	281	113	174	26	261	65	222
...
TOTAL (sup. rows)	12551	1313	2953	774	584	1509	1639	285	1924	19684	6104	15504	5229	16379	16087	5521

Table 4. Input sample for CA (total = 9 632 cells);
in white: active rows ($nrow = 543$) and active columns ($ncol = 8$);
in shades of grey: supplementary rows ($nrow = 59$) and supplementary columns
($ncol = 8$).

The script outputs the following result:

```
**Results of the Correspondence Analysis (CA)**  
Row variables have 543 categories, column variables have 8 categories  
The chi-square of independence between the two variables is equal to 24522.65  
(p-value = 0).
```

χ^2 has a very high value (24 522.65) and it is associated with the smallest possible p -value, implying that there is a relationship between the row variables and the column variables. Yet, we should be wary of using the magnitude of the χ^2 value to quantify the effect of the correlation between variables, because this value depends on the sample size. The total sample size (43 216) is small relative to the number of active cells in the table ($543 \times 8 = 4\,344$). This means that the table contains many cells with small or null values. We are far from meeting the assumptions of the χ^2 test, which stipulate that 80% of the sample size should be greater than 5, and the remaining 20% greater than 1. However, because CA is exploratory, it can be applied to tables even when the conditions of validity of the χ^2 statistic are not met (Greenacre, 2007). Given that the p -value is 0, the significance of the deviation of the table from independence is undeniable. In other words, the choice of adjectives and the choice of active variables are globally interdependent in the dataset.

Central to CA is the concept of profile. To obtain the profile of a row, each row frequency is divided by its row total. The average row profile is the average profile of all construction tokens put together. In a similar fashion, one obtains the profile of a column by dividing each column frequency by the column total. The average column profile is the average profile of all adjectives put together. Distances between profiles are measured with inertia (Φ^2), i.e. “the weighted average of squared χ^2 -distances between the row profiles and their average profile (similarly, between the column profiles and their average)” (Greenacre, 2007: 32). More concretely, CA interprets inertia geometrically to assess how far row/column profiles are from their respective average profiles. The total inertia of a table is the χ^2 statistic divided by the sample size. It is with inertia that CA measures how much variance there is in the table. In our data set, Φ^2 has a value of 1.13, which is relatively high. Therefore, we can expect data points to be more spread out on the map than if Φ^2 were lower. This is because there are noticeable differences between the data profiles.

Figure 1 maps the first two dimensions of the data. Dimension 1 is represented by the horizontal axis, and dimension 2 by the vertical axis. 8 adjectives have been plotted for illustrative purposes, as will appear below. The adjectives *expensive* and *pleasant* (in blue) are not associated with any particular construction and are very close to the average profiles (i.e. where the horizontal and vertical axes intersect). Given that *<quite ADJ>* is relatively close to the average column profile, it has a more representative profile than the other constructions.

The syntactic profiles of *quite* have negative coordinates and cluster on the left-hand side of the plot, whereas the syntactic profiles of *rather* have positive coordinates and cluster on the right-hand side. This means that, there is a clear divide between the syntactic profiles of *quite* and *rather* (in black) along the first dimension.

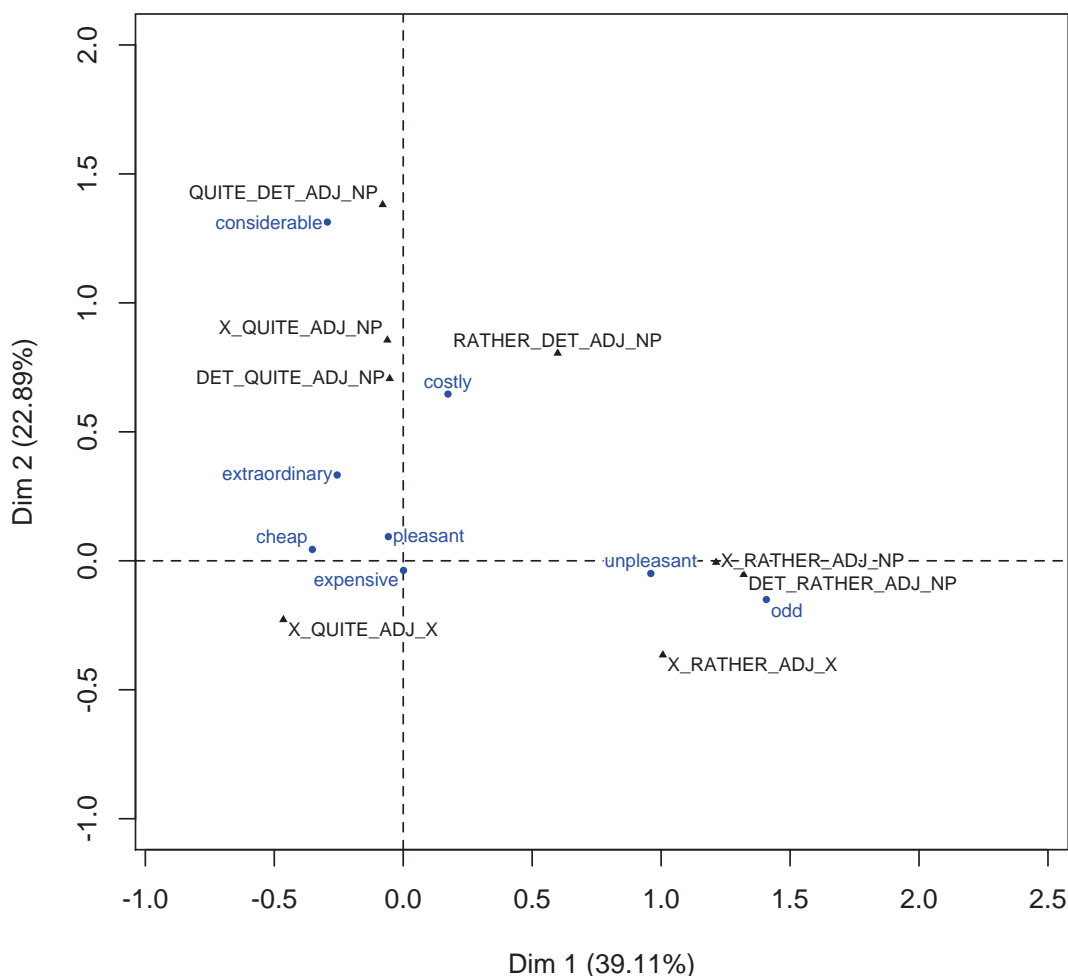


Figure 1. CA plot: a simultaneous representation of all active columns and 8 active rows.

Dimension 2 shows that the patterns of *quite* and *rather* subdivide into three levels in parallel fashion (Table 5). This parallel becomes clearer if we abstract away from dimension 1 (i.e. the horizontal axis).

	<i>quite</i>	<i>rather</i>
level 1	<quite ADJ>	<rather ADJ>
level 2	<DET quite ADJ NP>, <quite ADJ NP>	<DET rather ADJ NP>, <rather ADJ NP>
level 3	<quite DET ADJ NP>	<rather DET ADJ NP>

Table 5. Column clusters in dimension 2.

The relative distance to the average column profile is what determines each level (the higher the level index, the larger the distance from average).

To some extent *quite* and *rather* follow a functional division of labor. For example, *pleasant* co-occurs with <quite ADJ>, whereas its antonym *unpleasant* co-occurs with <rather ADJ NP>, <DET rather ADJ NP>, and <rather ADJ>. The same

three patterns attract *odd*, which connotes strangeness negatively, whereas <DET *quite* ADJ NP> and <*quite* ADJ> attract *extraordinary*, which connotes strangeness positively. Next, <DET *quite* ADJ NP> prefers *costly* over its less formal equivalent *expensive*, a sign that the choice of a syntactic pattern may be influenced by register. Little can be said about the proximity of <*quite* DET ADJ NP> and *considerable* besides a preference for this pattern to intensify adjectives that denote large proportions. If we projected more adjectives, we would see that *considerable* clusters with near-synonyms (*compelling, massive, extensive, big, large, etc.*), which would corroborate our claim.

In CA, supplementary rows and columns can also be plotted to help interpret the active rows and columns. As opposed to active elements (= white rows and columns in Table 4), supplementary elements do not contribute to the construction of the dimensions. They can still be positioned on the map, but only after Φ^2 has been calculated with the active elements. Given the number of rows in our dataset, plotting all 543 adjectives would make little sense, as the myriad of points would clutter the graph. To make these data points invisible while still being able to inspect their general profiles, the supplementary rows are projected instead (i.e. the rows in grey in Table 4). These rows group the 543 adjectives into the 59 semantic classes described above. Each row cell is the sum of the frequencies of all construction tokens whose adjectives illustrate the semantic class. Because of their intermediate granularity, the supplementary rows are easier to interpret than the active rows.

8 supplementary columns (i.e. the columns in grey in Table 4) are also projected to: (a) see if the relationship between the syntax of the intensifiers and the syntax of the adjectives brings new insights on our dataset, (b) see if the mode (*spoken* vs. *written*) plays a role in the choice of a construction over another, and (c) summarize the general profiles of *quite* and *rather* to evaluate how close or distant they are. Figure 3 provides a superimposed representation of supplementary rows (in blue) and supplementary columns (in black).

As expected, Figure 2 confirms most of the tendencies revealed in Figure 1. The position of the supplementary column variables on the first dimension corroborates the broad divide between *quite* and *rather* – *quite* constructions clustering to the left, and *rather* constructions to the right. Because *quite* constructions are closer to the average column profile than *rather* constructions, the profile of *rather* stands out in our dataset. As will appear below, the choice of *rather* over *quite* does not depend on the mode (since both *spoken* and *written* are close to the average column profile) but on the meaning of the intensified adjective.

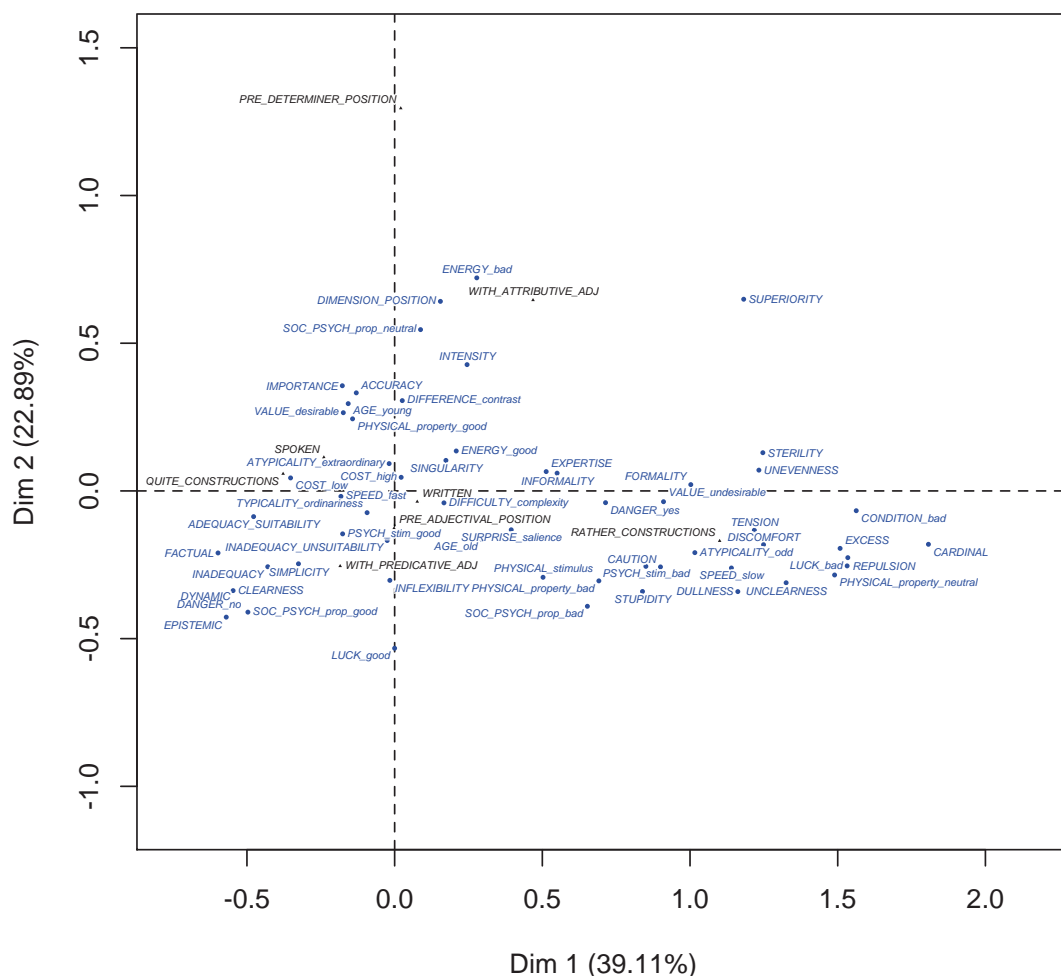


Figure 2. CA plot: a simultaneous representation of supplementary columns and supplementary rows.

Dimension 2 confirms the gradient of typicality inferred from Table 5. Typically, the intensifiers occur in pre-adjectival position, and the adjectives in the predicative position. Since *quite* and *rather* can only occur in pre-determiner position when the adjective is attributive, we have reasons to expect a strong correlation between these two variables. Such is not the case as the attributive variable is also seemingly attracted to the pre-adjectival position (bottom left part of the plot). When *quite* and *rather* occur in pre-determiner position, we find no significant proximity with any of the semantic classes of adjectives. The extreme position of the pre-determiner variable on the vertical axis makes this syntactic pattern stand out in our dataset. This is due to the overrepresentation of *<quite DET ADJ NP>* in the pre-determiner context (see Figure 1).

Regarding the semantic properties of *quite* and *rather* constructions, three configurations emerge. In the first configuration, each intensifier has an area of semantic specialization. In the second configuration, some meanings can be intensified

indiscriminately by *quite* or *rather*. In the third configuration, *quite* and *rather* divide up the task of intensifying two complementary aspects within a single conceptual domain.

In the first configuration, *quite* intensifies adjectives that denote:

- modal meanings (factual, dynamic, epistemic): *correct, false, true, able, aware, certain, likely*, etc.
- adequacy and inadequacy: *legitimate, acceptable, fair, unfair, inappropriate, unsatisfactory, unsuitable*, etc.
- cost (low or high): *cheap, expensive, costly*
- simplicity and complexity: *easy, simple, difficult, complex*, etc.
- typicality: *normal, ordinary, usual, familiar*, etc.

Rather intensifies adjectives that denote:

- formality and informality: *formal, strict, casual, informal*, etc.
- physical properties: *ugly, fat, stout, fit, nasal*, etc.
- physical stimuli: *quiet, loud, silent, sweet, bitter*, etc.
- meanings that are connoted negatively: stupidity, dullness, unclearness, unevenness, sterility, bad condition, excess, repulsion: *dumb, dull, cloudy, unequal, barren, battered, excessive, gruesome*, etc.

In the second configuration, *quite* and *rather* intensify the following meanings indiscriminately:

- energy (connoted negatively): *ferocious, violent*
- dimension/position: *large, little, long, short, high, low*, etc.
- social/psychological properties (connoted negatively): *desperate, shameful, busy, emotive, half-hearted, pessimistic, sad*.

The above meanings correlate with the choice of an adjective in attributive position, especially in the patterns <*quite* ADJ NP>, <DET *quite* ADJ NP>, and <*rather* DET ADJ NP> (see Figure 2)^v.

Table 6 illustrates the third configuration: when two connotations are available for a given conceptual domain, *rather* tends to select the negative alternative.

<i>quite</i>	<i>rather</i>
good luck (<i>lucky</i>)	bad luck (<i>unfortunate</i>)
atypicality (positive) (<i>outstanding, breathtaking</i>)	atypicality (negative) (<i>bizarre, unorthodox, puzzling</i>)
psychological stimuli (positive) (<i>exciting, interesting, moving</i>)	psychological stimuli (negative) (<i>disturbing, worrying, confusing</i>)
absence of danger (<i>harmless, safe</i>)	presence of danger (<i>threatening, dangerous</i>)
high speed (<i>quick, rapid</i>)	low speed (<i>slow, slower</i>)
desirable value (<i>good, neat, nice, perfect</i>)	undesirable value (<i>bad, poor, negative, nasty</i>)

Table 6. *The functional division of labor between quite and rather.*

All in all, the most typical profile in the dataset involves a construction in which:

- *quite* occurs in pre-adjectival position;
- the adjective occurs in predicative position;
- the adjective denotes a meaning with either neutral or positive connotations.

Incidentally, no striking relationship emerges between *rather* constructions and any of the syntactic properties of both intensifiers and adjectives. Despite its relative isolation, the attributive variable is not as atypical as the pre-determiner variable.

Apart from showing that, contrary to popular belief, the pre-determiner pattern is not significantly associated with written language (and the formality traditionally associated with it), text mode does not seem to play any discriminating role in the light of the variables ‘spoken’ and ‘written’. A more relevant option is to apprehend stylistic variation through more detailed modalities (fiction, news, conversation, etc.). Such information is available separately from the BYU-BNC for each query result, but given the number of tokens involved, manual extraction is not a viable option.

One interesting feature of exploratory methods is that one can look at the same dataset from different angles to make more tendencies appear. To this aim, one can even ignore irrelevant variables or add some variables one believes will prove relevant. To see how the pre-adjectival vs. pre-determiner alternation behaves in the light of more detailed contextual information, I imported contextual tags of the XML edition of the British National Corpus and augmented the dataset with three levels of contextual information instead of one for each observation: text mode, text type, and information regarding each text.

To remain within the strict limits of the pre-adjectival vs. pre-determiner alternation, I kept only those patterns that involve both alternants unquestionably, namely $\langle a(n) \text{ quite/rather ADJ NP} \rangle$ and $\langle \text{quite/rather } a(n) \text{ ADJ NP} \rangle$. Following the same logic, I ignored the cases where the adjective was in predicative position. Finally, the 6 variables corresponding to the remaining syntactic patterns listed in Table 1 were gathered into only two levels (‘pre-adjectival’ and ‘pre-determiner’) within a unique variable (‘construction’).

I obtained a dataset of 3 086 observations sampled in Table 7. This dataset consists of five categorical variables:

- construction: 2 modalities (pre-determiner and pre-adjectival),
- intensifier: 2 modalities (*quite* and *rather*),
- text mode: 2 modalities (written and spoken),
- text type: 8 modalities (academic writing, non-academic writing, fiction, news, other published writing, unpublished writing, conversation, and other spoken),
- text information: 66 modalities (e.g. S parliament, S interview, W pop lore, W newsp brdsht nat: editorial, etc.),
- semantic classes: 59 modalities (e.g. adequacy, accuracy_suitability, difficulty_complexity, etc.).

The dataset was submitted to multiple correspondence analysis, henceforth MCA (Greenacre & Blasius 2006; Le Roux 2010). As its name indicates, MCA is an extension of CA. It is used to analyze patterns of relationship between nominal variables. The general principles of CA outlined above also apply to MCA.

construction	intensifier	text_mode	text_type	text_info	sem_class
PREDETERMINER	QUITE	SPOKEN	OTHERSP	S pub debate	psych_stim_good
PREADJECTIVAL	QUITE	WRITTEN	FICTION	W fict prose	factual
PREADJECTIVAL	RATHER	WRITTEN	FICTION	W fict prose	dullness
PREADJECTIVAL	RATHER	WRITTEN	FICTION	W fict prose	atypicality_odd
PREADJECTIVAL	QUITE	SPOKEN	OTHERSP	S meeting	importance
PREADJECTIVAL	RATHER	WRITTEN	NONAC	W religion	difficulty_complexity
PREADJECTIVAL	RATHER	WRITTEN	NEWS	W newsp other: social	singularity
PREADJECTIVAL	QUITE	WRITTEN	NONAC	W biography	factual
PREDETERMINER	QUITE	SPOKEN	OTHERSP	S lect soc science	age_young
PREADJECTIVAL	RATHER	WRITTEN	NONAC	W nonAc: nat science	simplicity

Table 7. A sample from the second extraction, with contextual information

First, all the columns were treated as active except for the variable ‘text_info’. The variable ‘sem_class’ was ignored. Figure 3 maps the first two dimensions of the data. The plot differs from what Figures 1 and 2 show. In dimensions 1 and 2, there is a strong correlation between (a) *quite* and the pre-determiner construction in the upper left corner, and (b) between *rather* and the pre-adjectival construction in the lower right corner on the other. Along the horizontal axis (dimension 1), pre-determiner *quite* is correlated with the ‘spoken’ modality of the ‘text_mode’ variable, and more specifically with the two modalities of the ‘text_type’ variable (i.e. other spoken and conversation). In contrast, along the same axis, pre-adjectival *rather* is correlated with the ‘written’ modality of ‘text_mode’ and all the modalities of ‘text_type’ (i.e. academic writing, non-academic writing, fiction, news, other published writing, and unpublished writing). We should be wary of concluding that pre-determiner *quite* prefers spoken contexts and pre-adjectival *quite* prefers written contexts. Indeed, along the vertical axis (dimension 2), pre-determiner *quite* is found in one spoken context (conversation) and also four written contexts (other published writing, fiction, unpublished writing, and news). Along the same axis, pre-adjectival *rather* is found in one written contexts (academic writing) and also one spoken context (other spoken). Non-academic writing is specific to neither pre-determiner *quite* nor pre-adjectival *rather*.

‘epistemic’, or ‘physical_property_neutral’. Along the vertical axis, *quite* co-occurs with adjectives that have positive connotations: ‘value_desirable’ and ‘physical_property_good’. It also co-occurs with adjectives that denote neutral values such as spatial properties (‘dimension_position’) or social or psychological properties (‘soc_psych_prop_neutral’). Interestingly, in one context (age), *quite* is found with adjectives denoting opposite properties: ‘age_young’ and ‘age_old’. Along the same axis, *rather* co-occurs exclusively with adjectives that have negative connotations: ‘unclearness’, ‘atypicality_odd’, ‘dullness’, ‘value_undesirable’, ‘unevenness’, ‘epistemic’, ‘discomfort’, ‘singularity’, ‘physical_property_bad’, ‘repulsion’, and ‘luck_bad’. As seen above (see Table 6), when two connotations are available for a given conceptual domain, *rather* tends to select the negative alternative.

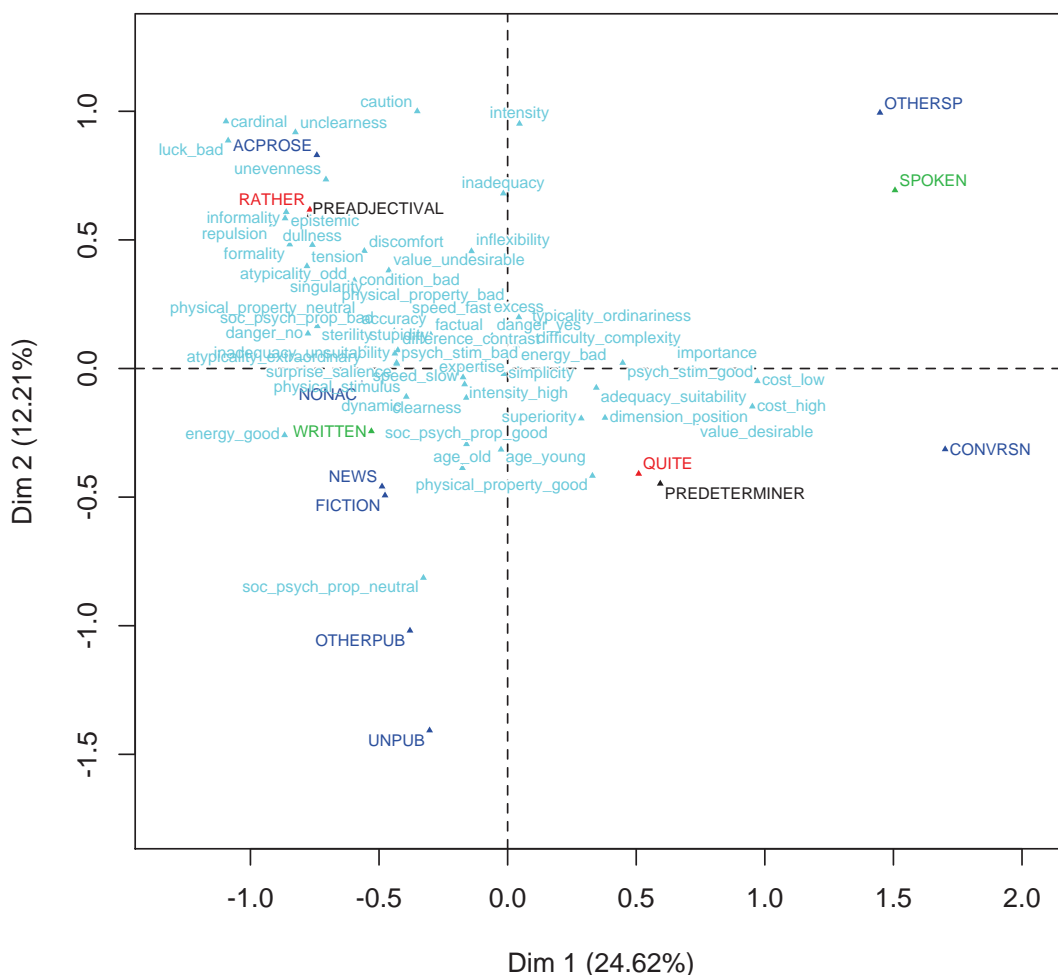


Figure 4. MCA plot : a simultaneous representation of pre-adjectival vs. pre-determiner constructions (active, in black), intensifiers (active, in red), text modes (active, in green), text types (active, in blue), and semantic classes (illustrative, in cyan)

Notwithstanding exceptions, the division of labor outlined above still holds. For example, pre-determiner *quite* is found with positively connoted modalities (e.g.

‘value_desirable’, ‘psych_stim_good’, or ‘physical_property_good’) whereas pre-adjectival *rather* is found with their negative counterparts (‘value_undesirable’, ‘psych_stim_bad’, ‘physical_property_bad’). Admittedly, not all the oppositions listed in Table 6 are found in the first two dimensions of the MCA plot in Figure 4 (for example, the ‘luck_good’ vs. ‘luck_bad’ is absent from this second dataset). It suggests that the behavior of the alternation differs slightly from the more general behavior of *quite* and *rather*.

Coming back to Allerton’s claim, we may concede that the choice of pre-determiner position over the default pre-adjectival position is sensitive to context or register. Yet, this choice is indeed more than a matter of style or formality both within the strict frame of the pre-determiner vs. pre-adjectival alternation, and beyond, i.e. in the broader frame of adjectival intensification involving *quite* and *rather*.

5. Discussion

I set out to capture usage-based relations between *quite* and *rather* from the perspective of their near synonymy. I argued that linguists should be wary of reducing the study of near synonymy to a matter of choices between lexemes only because we assumed that just as lexical choices reflect construals, so do their constructional expressions.

In the two CA plots (Figures 1 and 2), the differences between *quite* and *rather* are summarized graphically by their respective projections on the map: the constructional features of *quite* cluster on the left-hand side whereas the constructional features of *rather* cluster on the right-hand side. Affinities with syntactic and semantic profiles can be assessed spatially. I believe the method presented above teases out successfully the subtle mix of similarities and differences that holds between two near synonyms.

Because CA is based simultaneously on collocation data and formal and semantic variables, it is a clear improvement upon previous lexeme-based collocation studies on intensifiers that treat form, meaning, and context separately. Under the assumption that syntactic structure is an index of conceptual structure, the study of *quite* and *rather* gains more by correlating form and meaning than by relying on large, separate tables, each table illustrating one variable.

On the one hand, given that natural languages avoid true synonymy (Cruse 1986: 270), we should not be too surprised to see that *quite* and *rather* differ in some respect (cf. the first configuration in the previous section):

- (16) *It has been said that ‘the problem of turbulence has been solved’. Aside from the question of what is ‘the problem of turbulence’ (there are many), this can give a quite? rather false impression.* (BNC-J12)
- (17) *About 8 p.m. that evening Taff and I sat at the side of our slit trench eating what I thought was a rather? quite dubious lump of meat. Taff said it was lamb.* (BNC-A61)

On the other hand, given that speakers generalize over recurring experiences of language use, similar linguistic information can be stored at multiple levels in the inventory of linguistic signs. Distinct linguistic units that are not necessarily synonyms may thus share identical semantic content. Therefore, that *quite* and *rather* share substantial common ground should come as no surprise either (cf. the second configuration):

- (18) (...) *taboos persist and elicit sometimes quite violent reactions if they are broken (...)* (BNC-ECY)
- (19) *'Often,' he notes, 'we resembled a rather violent community welfare body rather than a group of revolutionaries.'* (BNC-APP)

The most remarkable result to emerge from the data is that when *quite* and *rather* operate within the same conceptual domain, they often impose complementary focal adjustments on the representation that the conceptual domain evokes, following what I called a division of labor (third configuration, Table 6). When adjectives have a negative connotation, they tend to occur in *rather* constructions:

- (20) *Three-tiered walls and arcades of massive pointed arches soar upwards in a quite breathtaking fashion.* (BNC-B3K)
- (21) *They met towards the end of January, and he kept the news from Theo until April. When he did let it out, he did so in a rather peculiar fashion, linking it to a quarrel with Mauve and casting it in a dramatic mode, with himself in the first and then third person.* (BNC-CBN)

Equally remarkable is the fact that, for a given pair of near synonyms, one construction can be more typical than the other. I defined typicality in terms of proximity relative to the average row/column profiles (where the first two dimensional axes intersect in Figure 3). In this respect, *quite* constructions are more typical than *rather* constructions, and within the family of *quite* constructions, those where the adverb occurs after the determiner and before a predicative adjective are more typical than the other configurations. At this stage, however, we should sound a voice of caution against an absolute interpretation of typicality. For example, the results showed that the choice of an attributive adjective was less typical than the choice of a predicative adjective, based on dimensions 1 and 2 in Figure 3^{vi}. Yet, at a finer grained level of analysis, we could also interpret the choice of an attributive adjective as more typical than the choice of a predicative adjective in the intensification of the following meanings: negatively connoted energy, social/psychological properties, dimension/position, intensity, and superiority. Because typicality is relative and constructional profiles are context-dependent, *quite* constructions are not always more typical than *rather* constructions.

In contrast to confirmatory statistical models, which provide accuracy scores of explanatory power, CA and MCA are exploratory: these methods merely tell us that the distributions in our dataset are statistically significant and not due to chance. Therefore, we should not infer any causal link between the variables.

Concerning CA specifically, it could be argued that the dataset is a sparse matrix (containing many cells with small or null values), as mentioned above. Although this could be a problem for confirmatory statistics, the fact that the conditions of validity of the χ^2 statistics are not met does not much affect exploratory methods. Indeed, CA does not so much prove the existence of a relationship between columns and rows as the capacity of a limited set of data to make that relationship visible (Husson et al. 2011).

Finally, one anonymous reviewer considers that *quite* and *rather* are genuine intensifiers only when these adverbs express high degree, as opposed to when they express a moderate degree, as in *well, she's quite beautiful... but not quite* (the reviewed

provided that example). The same reviewer wishes a distinction had been made between these two degrees. Following the literature on intensifiers, which already provides a response to this critique, I consider that *quite* and *rather* are intensifiers as long as they index an adjectival property on a scale, regardless of the specification of degree. As degree words, *quite* and *rather* give specifications of degree regarding the adjectives that they modify (Bolinger 1972). When *quite* and *rather* set the qualities that gradable adjectives denote to a moderate level, they function as ‘compromisers’ (Paradis 1994), ‘moderators’ (Paradis 1997), or ‘downtoners’ (Nevalainen & Rissanen 2002), along with adverbs such as *fairly*, *mildly*, *moderately*, *partially*, *pretty*, *relatively*, etc. When *quite* and *rather* set gradable adjectival qualities to a high level, along with *greatly*, *really*, *so*, *strongly*, *too*, *very*, etc., they function as ‘boosters’ (Paradis 1997). Determining whether degree modifiers are moderators, boosters, or maximizers is indeed relevant, but it cannot be done using only corpus-based techniques, mainly because semantic context is not always helpful. This has to be done experimentally, as exemplified by Paradis (1994, 1997), who interprets <intensifier + ADJ> collocations in combination with intonation patterns in the prosodically annotated London-Lund Corpus of Spoken English.

Despite the above caveats, and pending confirmatory statistics and experimental verification, I believe that my approach contributes some new understanding of intensifiers (within the frame of the pre-determiner vs. pre-adjectival alternation and beyond), constructions, and collocation analysis. Once distinctive collexemes have been identified, the data can be used again safely in multi-way tables to map correlations between intensifiers, adjectives, and information concerning context of use and constructional idiosyncrasies.

6. Conclusion

I have combined several statistical methods to compare the constructional profiles of *quite* and *rather*, taking into account both semantic and syntactic factors. First, I asserted the need for better statistics in the collocation-based study of intensifiers. Regardless of the specific quantitative method that one adopts, it is particularly important to filter away co-occurring pairs that are unrealistically too frequent or too rare. It is also important to compute collocation strengths with statistics that do not violate the distributional assumptions that are specific to language data. Next, I augmented the collocation data with syntactic and semantic variables that license the use of *quite* and *rather* in context. Finally, I used two multifactorial methods to explore and visualize two datasets: one illustrating the differences between *quite* and *rather*, another one focusing on the pre-determiner vs. pre-adjectival alternation. Each time, similarities and differences were synthesized in one graph.

This has led me to assess the near synonymy of not just two lexemes (*quite* and *rather*), but a network of construction types and subtypes. All assume different profiles depending on what conceptual domains (or what parts of a specific conceptual domain) intensification bears on. Through the choice of a syntactic profile, the speaker adopts a focal adjustment on the meaning of an adjectival property and influences the way that the hearer interprets this conceptual representation in context.

Based on a combination of MDCA and CA, results show that *quite* and *rather* constructions are near-synonyms. On the one hand, both subdivide into three syntactic levels following a gradient of prototypicality, from pre-adjectival patterns to pre-determiner patterns. Both also share enough semantic properties to intensify adjectives

with similar meanings (e.g. negatively connoted energy and social/psychological properties, and dimension/position). On the other hand, *quite* and *rather* constructions are distinct enough to divide up the task of intensifying certain categories of adjectival meanings. The most striking tendency is that *quite* constructions have a preference for the intensification of positive meanings (e.g. *quick, harmless, lucky*, etc.), and *rather* constructions have a preference for the intensification of negative meanings (e.g. *slow, threatening, unfortunate*, etc.).

Based on a combination of MDCA and MCA, results also show that *quite* displays a neat preference for the pre-determiner pattern, and *rather* for the pre-adjectival pattern in the strict frame of the pre-determiner vs. pre-adjectival alternation. Finally, although relevant to some extent, context does not fully account for the pre-determiner vs. pre-adjectival alternation. The abovementioned division of labor plays a central role in the choice of one alternant over the other.

The approach described in this paper can be extended to other intensifiers and other forms of intensification in English. It can also be extended to any linguistic paradigm, providing one wants to show how the paradigm is structured, for instance through the description of entrenchment continua between typical profiles. This is immediately relevant to those linguists who subscribe to an inventory approach to the mental representation of grammar, as formulated in Goldberg (2006, 2009) and Cognitive Construction Grammar in general (Langacker 2009). Cognitive Construction Grammar posits that linguistic knowledge consists of a network of families of constructions, and advocates a usage-based model to explain how constructions are stored in the taxonomies. This paper should therefore be seen as a quantitative corpus-based contribution to a usage-based description of constructional networks.

Appendix

SEMANTIC_CLASSES (as coded)	ADJECTIVES (examples)
ACCURACY	<i>accurate, subtle</i>
ADEQUACY_SUITABILITY	<i>legitimate, adequate</i>
AGE_old	<i>archaic, old</i>
AGE_young	<i>early, junior, new</i>
ATYPICALITY_extraordinary	<i>extreme, amazing</i>
ATYPICALITY_odd	<i>bizarre, curious</i>
CARDINAL	<i>later, then</i>
CAUTION	<i>careful, cautious</i>
CLEARNESS	<i>clear, evident</i>
CONDITION_bad	<i>untidy, battered</i>
COST_high	<i>expensive, costly</i>
COST_low	<i>cheap</i>
DANGER_no	<i>harmless, safe</i>
DANGER_yes	<i>threatening, dangerous</i>
DIFFERENCE_contrast	<i>contrary, contrasting</i>
DIFFICULTY_complexity	<i>complex, complicated</i>
DIMENSION_POSITION	<i>big, broad, high</i>
DISCOMFORT	<i>clismy, oppressive</i>
DULLNESS	<i>colourless, dull, drab</i>

DYNAMIC	<i>active, effective</i>
ENERGY_bad	<i>ferocious, violent</i>
EPISTEMIC	<i>certain, doubtful</i>
EXCESS	<i>exaggerated, excessive</i>
EXPERTISE	<i>academic, specialised</i>
FACTUAL	<i>correct, false</i>
FORMALITY	<i>formal, strict</i>
IMPORTANCE	<i>fundamental, important</i>
INADEQUACY	<i>useless, inadequate</i>
INADEQUACY_UNSUITABILITY	<i>inappropriate, unsuitable</i>
INFLEXIBILITY	<i>firm, forbidding</i>
INFORMALITY	<i>casual</i>
INFORMALITY	<i>informal</i>
INTENSITY_high	<i>abrupt, radical</i>
LUCK_bad	<i>unfortunate</i>
LUCK_good	<i>lucky</i>
PHYSICAL_property_bad	<i>ugly, smelly</i>
PHYSICAL_property_good	<i>stout, hefty</i>
PHYSICAL_property_neutral	<i>nasal, static</i>
PHYSICAL_stimulus	<i>sweet, bitter</i>
PSYCH_stim_bad	<i>alarming, depressing</i>
PSYCH_stim_good	<i>amusing, exciting</i>
REPULSION	<i>crude, gruesome</i>
SIMPLICITY	<i>easier, easy</i>
SINGULARITY	<i>idiosyncratic, distinctive</i>
SOC_PSYCH_prop_bad	<i>unnerving, upset</i>
SOC_PSYCH_prop_good	<i>fascinating, honest</i>
SOC_PSYCH_prop_neutral	<i>human, personal</i>
SPEED_fast	<i>quick, rapid</i>
SPEED_slow	<i>slow, slower</i>
STERILITY	<i>sterile, barren</i>
STUPIDITY	<i>foolish, absurd, dumb</i>
SUPERIORITY	<i>leading, superior</i>
SURPRISE_salience	<i>stunning, marked</i>
TENSION	<i>strained, tense</i>
TYPICALITY_ordinariness	<i>familiar, ordinary</i>
UNCLEARNESS	<i>abstract, cloudy</i>
UNEVENNESS	<i>unequal, erratic</i>
VALUE_desirable	<i>superb, advanced</i>
VALUE_undesirable	<i>mixed, mediocre</i>

References

- Allerton, D. J. "English Intensifiers and Their Idiosyncrasies." *Language Topics: Essays in Honour of Michael Halliday*. Eds. Steele, Ross and Terry Threadgold. Vol. 2. Amsterdam: John Benjamins, 1987. 15-31.
- Altenberg, Bengt. "Amplifier Collocations in Spoken English." *English Computer Corpora: Selected Papers and Research Guide*. Eds. Johansson, Stig and Anna-Brita Stenström. Vol. 3. Topics in English Linguistics. Berlin; New York: Mouton de Gruyter, 1991. 127-49.
- Benzécri, Jean-Paul. *Analyse des correspondances, exposé élémentaire*. Pratique De L'analyse Des Données. Vol. 1. Paris : Dunod, 1984.
- . *L'analyse des données*. Vol. 2 L'analyse des correspondances. Paris Bruxelles Montréal : Dunod, 1973.
- Bolinger, Dwight. *Meaning and Form*. English Language Series No 11. London; New York: Longman, 1977.
- Bolinger, Dwight Le Merton. *Degree Words*. Janua Linguarum. Series Maior. Vol. 53. The Hague: Mouton, 1972.
- Burnard, Lou. "Reference Guide for the British National Corpus (World Edition)." 2000. Web.
- Bybee, Joan L. "From Usage to Grammar: The Mind's Response to Repetition." *Language* 82.4 (2006): 711-33.
- Clark, Eve V. "The Principle of Contrast: A Constraint on Language Acquisition." *Mechanims of Language Acquisition*. Ed. MacWhinney, Brian. Hillsdale, N.J.: Lawrence Erlbaum Associates, 1987. 1-33.
- Cruse, D. A. *Lexical Semantics*. Cambridge Textbooks in Linguistics. Cambridge Cambridgeshire ; New York: Cambridge University Press, 1986.
- Davies, Mark. "Byu-Bnc. Based on the British National Corpus from Oxford University Press. Available Online at [Http://Corpus.Byu.Edu/Bnc](http://Corpus.Byu.Edu/Bnc)." 2004. Web.
- Desagulier, Guillaume. "Le statut de la fréquence dans les grammaires de constructions : *simple comme bonjour?*" *Langages* 197.1 (2015): 99-128.
- . "Visualizing Distances in a Set of near Synonyms: *rather, quite, fairly, and pretty.*" *Corpus Methods for Semantics: Quantitative Studies in Polysemy and Synonymy*. Eds. Glynn, Dylan and Justyna Robinson. Amsterdam: John Benjamins, 2014.
- Divjak, Dagmar. *Structuring the Lexicon: A Clustered Model for Near-Synonymy*. Berlin; New York: Mouton de Gruyter, 2010.
- . "Ways of Intending: Delineating and Structuring near-Synonyms." *Corpora in Cognitive Linguistics*. Eds. Gries, Stefan and Anatol Stefanowitsch. Vol. 172. Trends in Linguistics. Studies and Monographs. Berlin; New York: Mouton de Gruyter, 2006. 19-56.
- Dixon, Robert M. W., and Alexandra Y. Aikhenvald. *Adjective Classes: A Cross-Linguistic Typology*. Oxford: Oxford University Press, 2004.
- Geeraerts, Dirk. *Theories of Lexical Semantics*. Oxford; New York: Oxford University Press, 2010.
- Gibbs, Raymond W. "Why Cognitive Linguists Should Care More About Empirical Methods." *Methods in Cognitive Linguistics*. Eds. Gonzalez-Marquez, Monica, et al. Amsterdam: John Benjamins, 2007. 2-18.
- Gilbert, Eric. "Quite, Rather." *Cahiers De Recherche, Grammaire Anglaise*. Vol. 4. Gap: Ophrys, 1989. 4-61.

- Gilquin, Gaetanelle. "The Verb Slot in Causative Constructions. Finding the Best Fit." *Constructions* (2007).
- Glynn, Dylan. "Correspondence Analysis: Exploring Data and Identifying Patterns." *Corpus Methods for Semantics: Quantitative Studies in Polysemy and Synonymy*. Eds. Glynn, Dylan and Justyna A. Robinson. Human Cognitive Processing. Amsterdam: John Benjamins, 2014. 443-85.
- Goldberg, Adele E. *Constructions : A Construction Grammar Approach to Argument Structure*. Cognitive Theory of Language and Culture. Chicago: University of Chicago Press, 1995.
- . *Constructions at Work : The Nature of Generalization in Language*. [Oxford Linguistics]. Oxford; New York: Oxford University Press, 2006.
- . "Constructions Work." *Cognitive Linguistics* 20.1 (2009): 201-24.
- Greenacre, Michael J. *Correspondence Analysis in Practice*. Interdisciplinary Statistics Series. Vol. 2. Boca Raton: Chapman & Hall/CRC, 2007.
- Greenacre, Michael J., and Jörg Blasius. *Multiple Correspondence Analysis and Related Methods*. Statistics in the Social and Behavioral Sciences Series. Boca Raton: Chapman & Hall/CRC, 2006.
- Gries, Stefan Thomas. "Coll.Analysis 3.2. A Program for R for Windows 2.X." 2007.
- Gries, Stefan Thomas, Beate Hampe, and Doris Schönefeld. "Converging Evidence: Bringing Together Experimental and Corpus Data on the Association of Verbs and Constructions." *Cognitive Linguistics* 16.4 (2005): 635-76.
- Gries, Stefan Thomas, and Anatol Stefanowitsch. "Co-Varying Collexemes in the *into-Causative*." *Language, Culture, and Mind*. Eds. Achard, Michel and Suzanne Kemmer. Stanford, Calif.: CSLI, 2004. 225-36.
- . "Extending Collostructional Analysis: A Corpus-Based Perspective on 'Alternations'." *International Journal of Corpus Linguistics* 9.1 (2004): 97-129.
- Hilpert, Martin. "Distinctive Collexeme Analysis and Diachrony." *Corpus Linguistics and Linguistic Theory* 2.2 (2006): 243-56.
- . *Germanic Future Constructions : A Usage-Based Approach to Language Change*. Constructional Approaches to Language. Amsterdam; Philadelphia: John Benjamins Pub. Co, 2008.
- Husson, Francois, Julie Josse, Jérôme Pagès, and Sébastien Lê. "FactoMineR, an R Package Dedicated for Multivariate Analysis." 2009.
- Husson, Francois, Sebastien Lê, and Jerome Pagès. *Exploratory Multivariate Analysis by Example Using R*. Chapman & Hall/Crc Computer Science and Data Analysis. Boca Raton: CRC Press, 2011.
- Kennedy, Graeme. "Amplifier Collocations in the British National Corpus: Implications for English Language Teaching." *TESOL Quarterly* 37.3 (2003): 467-87.
- Langacker, Ronald W. "Cognitive (Construction) Grammar." *Cognitive Linguistics* 20.1 (2009): 167-76.
- . *Cognitive Grammar : A Basic Introduction*. Oxford: Oxford University Press, 2008.
- . "A Dynamic Usage-Based Model." *Usage-Based Models of Language*. Eds. Barlow, Michael and Suzanne Kemmer. Stanford: CSLI Publications, 2000. 1-64.
- . *Foundations of Cognitive Grammar*. Vol. 1 (Theoretical Prerequisites). Stanford, CA: Stanford University Press, 1987.
- . *Foundations of Cognitive Grammar*. Vol. 2 (Descriptive application) vols: Stanford, 1991.

- . *Grammar and Conceptualization*. Cognitive Linguistics Research. Berlin: Mouton de Gruyter, 1999.
- . "A Usage-Based Model." *Topics in Cognitive Linguistics*. Ed. Rudzka-Ostyn, Brygida. Amsterdam ; Philadelphia: John Benjamins, 1988. 127-61.
- Le Roux, Brigitte. *Multiple Correspondence Analysis*. Quantitative Applications in the Social Sciences ; 163. Ed. Rouanet, Henry. London: SAGE, 2010.
- Lorenz, Gunter R. *Adjective Intensification: Learners Versus Native Speakers: A Corpus Study of Argumentative Writing*. Amsterdam: Rodopi, 1999.
- . "Really Worthwhile or Not Really Significant? A Corpus-Based Approach to the Delexicalization and Grammaticalization of Intensifiers in Modern English." *New Reflections on Grammaticalization*. Eds. Wischer, I. and G. Diewald. Amsterdam: John Benjamins, 2002. 143-61.
- Nevalainen, Terttu, and Matti Rissanen. "Fairly Pretty or Pretty Fair? On the Development and Grammaticalization of English Downtoners." *Language Sciences* 24.3-4 (2002): 359-80.
- Paradis, Carita. "Compromisers – a Notional Paradigm." *Hermes* 13 (1994): 157-67.
- . "Degree Modifiers of Adjectives in Spoken British English." Lund: Lund University Press, 1997.
- 'quite, adv.'. *OED Online*. December 2012. Oxford University Press. <https://www-oed-com>. (accessed January 23, 2013).
- R Core Team. "R: A Language and Environment for Statistical Computing." Vienna: Austria: R Foundation for Statistical Computing, 2015.
- 'rather, adv.'. *OED Online*. December 2012. Oxford University Press. <https://www-oed-com>. (accessed January 23, 2013).
- Simon-Vandenberg, Anne-Marie. "Almost Certainly and Most Definitely: Degree Modifiers and Epistemic Stance." *Special Issue: Pragmatic and Discourse-Analytic Approaches to Present-Day English* 40.9 (2008): 1521-42.
- Stefanowitsch, Anatol. "Collostructional Analysis." *The Oxford Handbook of Construction Grammar*. Eds. Hoffmann, Thomas and Graeme Trousdale. Oxford: Oxford University Press, 2013. 290-306.
- Stefanowitsch, Anatol, and Stefan Thomas Gries. "Collostructions: Investigating the Interaction of Words and Constructions." *International Journal of Corpus Linguistics* 8.2 (2003): 209-43.
- . "Covarying Collexemes." *Corpus Linguistics and Linguistic Theory* 1.1 (2005): 1-46.
- Stoffel, Cornelius. "Intensives and Down-Toners: A Study in English Adverbs." Heidelberg: Carl Winter, 1901.
- Tummers, José, Kris Heylen, and Dirk Geeraerts. "Usage-Based Approaches in Cognitive Linguistics: A Technical State of the Art." *Corpus Linguistics and Linguistic Theory* 1.2 (2005): 225-61.

ⁱ Throughout this paper, examples from the British National Corpus appear with an indication of their corpus file. Examples without an indication of source are made up.

ⁱⁱ For example, the adverbs *nearly* and *almost* both express near-completion on a path (both literally and figuratively), but they impose different ways of construing this conceptual content: *nearly* profiles near-completion from within the path, whereas *almost* profiles near-completion from its endpoint.

ⁱⁱⁱ Here, the acceptability judgment is valid for British English only. American English accepts *quite curious* and *rather perfect* more readily.

^{iv} These are only tendencies. The BNC contains examples where *formal* and *loud* have neutral interpretations.

^v The position of the semantic class ‘superiority’ is marginal. This is an effect of the low frequencies of its members (*leading*, *superior*).

^{vi} In CA, the choice of dimensions does not affect the shape of the cloud of data points but the viewpoint that we adopt on the cloud. If we observed the cloud from a different angle (i.e. along other dimensions) we could arrange the data points in such a way as to place variables that were atypical along the first two dimensions in a typical position relative to the average row/column profiles. Such a representation would have to be handled with due care because it would certainly not capture a representative amount of inertia, but it would confirm that the typicality of profiles should not be interpreted in the absolute.