

**QUANTIFYING COGNITIVE SYNONYMY**  
**A CORPUS STUDY OF ENGLISH MODERATORS**  
***RATHER, QUITE, FAIRLY & PRETTY***

Workshop "Empirical approaches to social cognition and emergent language structure"  
(convenors: D. Glynn and K. Krawczak).

Guillaume DESAGULIER

Université Paris 8 – St Denis  
UMR 7114 MoDyCo (Paris Ouest Nanterre La Défense – CNRS)

gdesagulier@univ-paris8.fr  
<http://www.univ-paris8.fr/desagulier/home>

## 1. INTRODUCTION

### i. NEAR-SYNONYMY, CONCEPTUAL CONTENT, AND CONSTRUAL

- In Cognitive Grammar terms, **synonymous expressions** have identical conceptual content and impose the same construal upon that conceptual content (Langacker, 1987; 1991; 2008);
- **near-synonyms** share the same conceptual content but differ in terms of construal.

☞ If we manage to measure and represent graphically similarities in conceptual content and differences in construal, we shall be able to observe fine semantic differences in sets of near-synonyms.

### ii. CASE STUDY: RATHER, QUITE, FAIRLY, & PRETTY

*Rather, quite, fairly, and pretty* set the qualities that gradable adjectives denote to a moderate level (i.e. they index the properties of the adjectives they modify as 'not fully X'). Along with *moderately* and *relatively*, these degree modifiers are known as 'moderators' (Paradis, 1997).

- investigation restricted to <ADV + ADJ>:

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

- adverbs used as word modifiers/constituents of AP (Stoffel, 1901);
- they scale inherent qualities or properties of the heads;
- **because these adverbs are near-synonyms, we may expect them to share identical conceptual content but differ as to how this conceptual content is construed.**

- entrenched types of the <moderator + adjective> construction provides access to a conceptual structure a non-compositional way (Goldberg, 1995; 2003; 2006; Langacker, 1987; 2008)

### iii. PREVIOUS RESEARCH

- Rather, quite, fairly, and pretty* are typologically challenging...
  - quite* Allerton (1987), Paradis (1997), Athanasiadou (2007)
- (5) *The market is quite big and quite old.* (COCA) -> moderator
- (6) *That was quite excellent.* (COCA) -> 'booster', i.e. high degree
- (7) *The film is quite good.* -> moderator or booster ?
  - rather* (Gilbert, 1989)
- (8) *She was quite large actually, but compared to Manny Fox she seemed rather frail* -> moderator
- (9) *rather unique* -> high degree
- ...and semantically hard to distinguish (Downing, 2006)
  - quite* "moderate but unequivocal intensification of the adjective";
  - rather* "understatement"; "mitigation"; "politeness";
  - pretty* "*quite* but not completely";
  - fairly* "an almost large or reasonable degree of a quality"
- Moderators are **cognitive synonyms** (Paradis, 1994: 160; 1997: 71) -> substituting one for the other does not change the truth-value of the proposition (Cruse, 1986: 270ff). Cognitive synonyms share identical conceptual content and differ only in style, register, and connotation. In other words, cognitive synonymy is a matter of both **sameness and difference**.
- Paradis (1997, 62) -> **the semantic relation between degree modifiers and adjectives is bidirectional** (the adjective selects a degree modifier, which in turn restricts the interpretation of the adjective)
- Previous corpus-based studies of degree modifiers: Altenberg (1991), Paradis (1997), Lorenz (2002), Kennedy (2003), Simon-Vandenberg (2008);
  - o Problem: inadequate measures of the two-way interaction between collocants, e.g. raw freq, %, counts per n-words, MI (Kennedy, 2003); see Kilgarriff (2001).

### iv. RESEARCH QUESTIONS

- How can we account for sameness and difference in context?
- How can we visualize near-synonymy?
- How can we account for bi-directional semantic relationship between moderators and adjectives?

#### Proposed solutions:

- we inspect co-occurrences of moderators and adjectives within constructional patterns because both moderators and adjectives contribute to the meaning of the <moderator + adjective> construction;
- we combine univariate and multivariate statistics to map usage patterns and conceptual structure.

## 2. METHODS

### i. CORPUS

COCA (Davies, 1990-present) – 415M words

Extraction: INTENSIFIER + ADJ (R1 & R2) + clean up + tabulations

### ii. STATISTICS

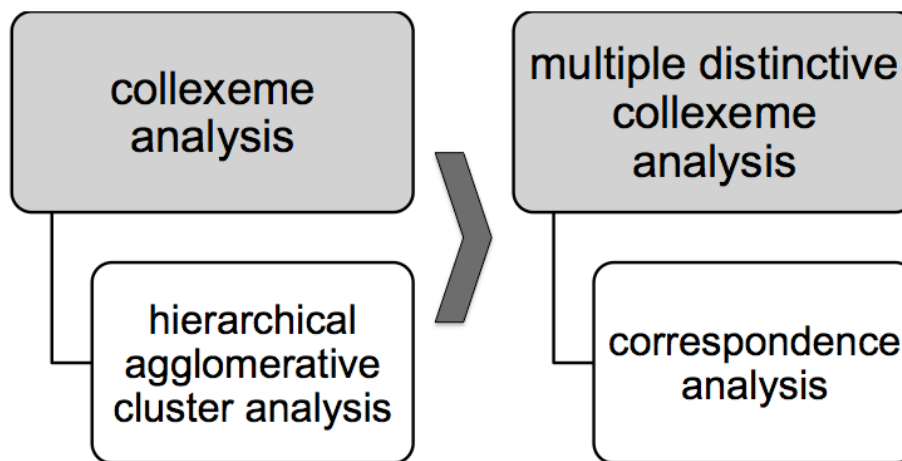


Figure 1. Collostructional analysis as input for two multifactorial methods

Collexeme analysis: Stefanowitsch & Gries (2003)

Multiple distinctive collexeme analysis (Gilquin, 2007; Gries & Stefanowitsch, 2004)

Hierarchical cluster analysis (Divjak, 2006; 2010; Divjak & Gries, 2008)

Correspondence analysis (Benzécri, 1973; Benzécri, 1984; Greenacre, 2007)

## 3. RESULTS

### i. COLLEXEME ANALYSIS

rather		quite		fairly		pretty	
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>straightforward</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

Table 1. Collexeme analysis of *rather*, *quite*, *fairly*, and *pretty*

ii. COLLEXEME ANALYSIS AS INPUT FOR HIERARCHICAL CLUSTER ANALYSIS

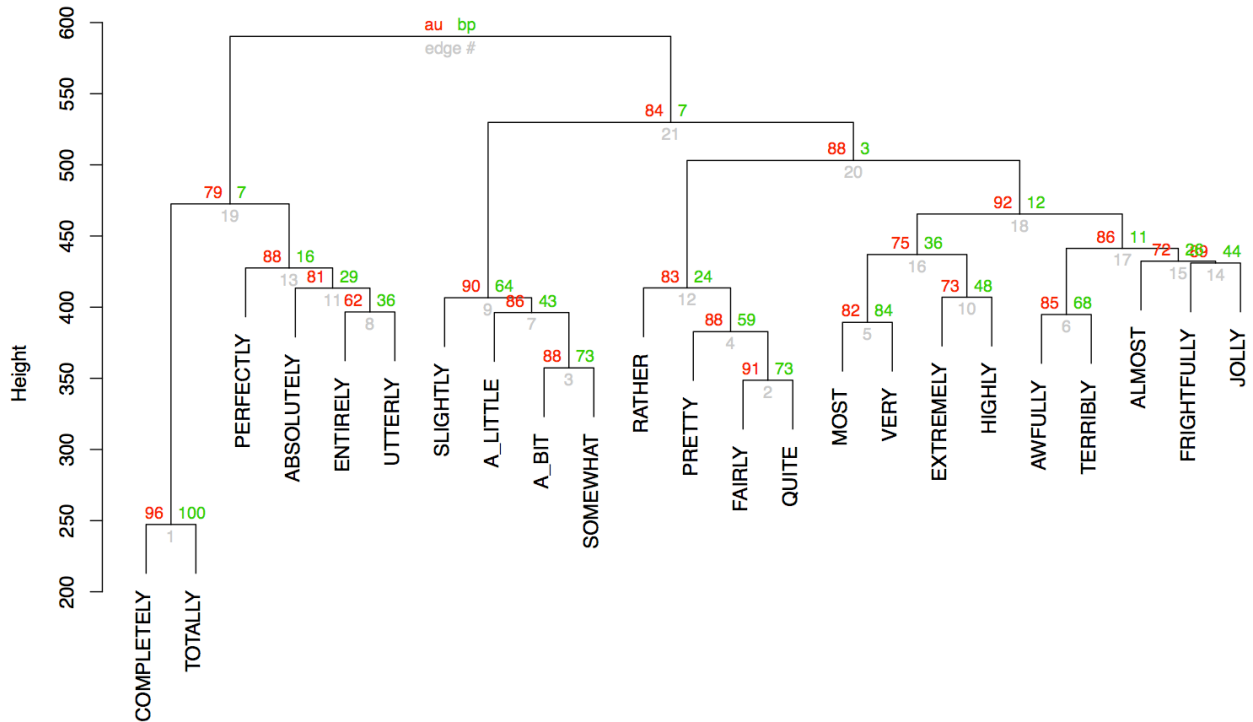


Figure 2. Cluster dendrogram of 23 degree modifiers of adjectives in English, clustered according to their adjectival collexemes (distance: Canberra; cluster method: ward)

iii. MULTIPLE DISTINCTIVE COLLEXEME ANALYSIS

<b>RATHER + ADJ</b>	obs freq	exp freq	pbin_RATHER	SumAbsDev
ODD	74	14.51	38.99	56.75
UNUSUAL	116	37.58	30.40	46.39
STRANGE	80	25.70	21.54	38.70
VAGUE	56	12.71	24.75	35.86
DIFFICULT	129	65.35	13.89	35.86
SIMPLISTIC	28	4.70	18.31	28.42
LENGTHY	36	9.26	13.81	27.53
PECULIAR	31	5.94	17.21	27.08
BIZARRE	51	14.09	17.46	25.83
CURIOUS	29	5.94	14.91	25.09
FORMAL	29	5.94	14.91	23.49

Table 2. MDCA of *rather*

<b>QUITE + ADJ</b>	obs freq	exp freq	pbin_QUITE	SumAbsDev
DIFFERENT	2247	825.30	Inf	Inf
RIGHT	548	184.75	238.72	410.01
POSSIBLE	458	150.95	216.96	374.49
READY	300	104.02	120.33	206.00
TRUE	235	91.23	70.29	120.24
LIKELY	206	76.46	69.12	118.13
SIMILAR	328	148.32	66.10	116.57
CAPABLE	152	51.19	66.87	114.97
WILLING	142	49.55	56.31	96.25
CORRECT	106	36.10	45.22	78.33

Table 3. MDCA of *quite*

<b>FAIRLY + ADJ</b>	obs freq	exp freq	pbin_FAIRLY	SumAbsDev
COMMON	278	83.86	76.85	145.26
EASY	337	108.56	83.94	130.39
NEW	202	43.38	88.58	123.43
CONSTANT	117	16.63	84.16	123.41
RECENT	122	20.36	72.50	109.92
CERTAIN	201	66.03	49.00	98.80
TYPICAL	154	35.55	61.99	89.24
CONSISTENT	130	33.26	45.98	68.47
STRAIGHTFORWARD	123	30.73	44.94	65.25
REGULAR	70	12.05	40.51	59.78

Table 4. MDCA of *fairly*

<b>PRETTY + ADJ</b>	obs freq	exp freq	pbin_PRETTY	SumAbsDev
GOOD	7731	3613.04	Inf	Inf
BAD	758	343.37	201.75	349.84
COOL	488	214.45	144.61	252.23
TOUGH	489	226.02	122.02	213.70
BIG	591	295.44	113.65	204.40
HARD	447	225.61	83.52	143.97
SCARY	212	98.34	52.78	93.43
SMART	196	91.32	48.17	83.66
AMAZING	347	195.03	44.89	82.05
CLOSE	498	314.03	40.52	77.60

Table 5. MDCA of *pretty*

iv. MDCA AS INPUT FOR CORRESPONDENCE ANALYSIS

ADJECTIVE (distinctive collexeme)	FAIRLY	PRETTY	QUITE	RATHER
ABLE	3	4	67	0
ABSTRACT	14	7	5	22
ACCURATE	83	66	109	4
AMAZING	6	347	97	22
AWARE	1	11	109	0
AWFUL	0	100	12	6
AWKWARD	0	12	7	27
BAD	19	758	32	22
BEAUTIFUL	1	2	129	21
BIG	56	591	45	23
...	...	...	...	...

Table 6. Input for correspondence analysis (sampled).

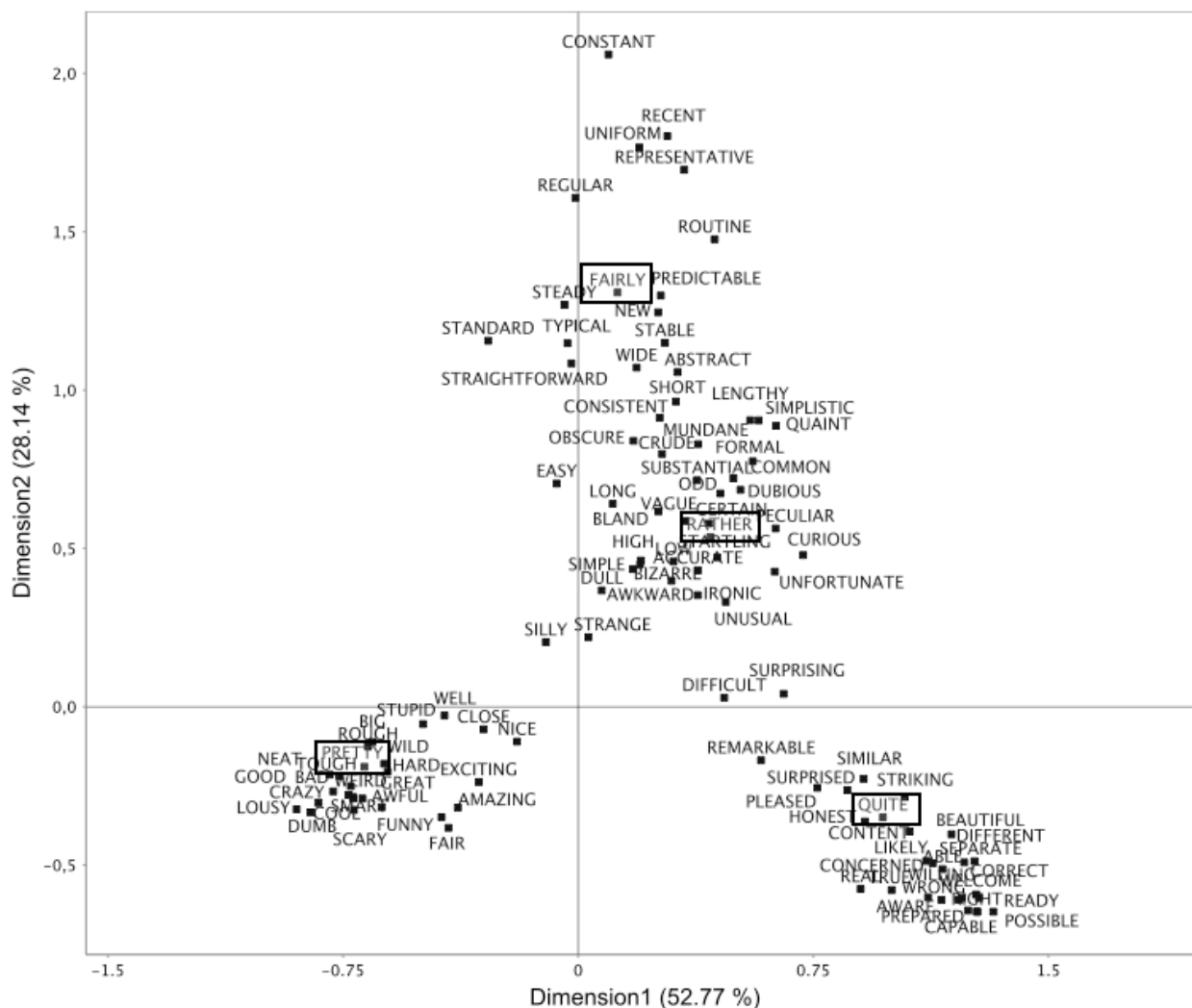


Figure 3. CA of the <moderator + adjective> construction in COCA (graphic output).

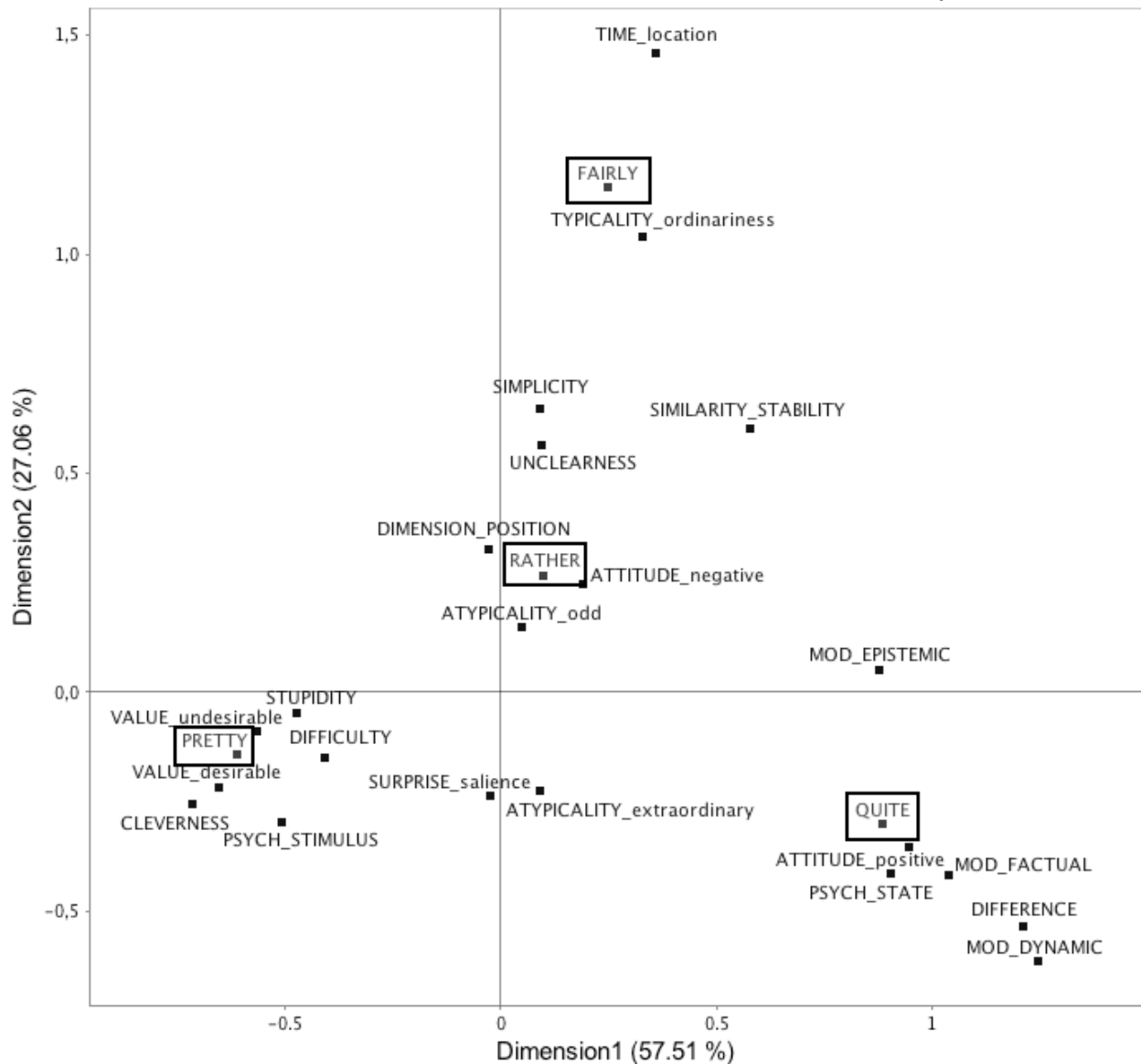


Figure 4. CA of the <moderator + adjective> construction in COCA (graphic output with semantic annotation).

#### 4. DISCUSSION

- Limits:

- No corpus, whatever its size is representative enough;
- Need for confirmatory stats & experimental verification;
- Aspects that have been left aside: syntactic idiosyncrasies (Allerton, 1987), grading force (Paradis, 1997), subjectivity (Nevalainen & Rissanen, 2002; Athanasiadou, 2007) - but the methodology proposed here can handle that.

- Assets:

- Once distinctive collexemes have been identified, raw frequencies can be used again safely in a multi-way table to map correlations between (a) moderators, (b) adjectives, (c) moderators and adjectives;
- One graph synthesizes similarities and differences;
- The CA plot makes it possible to determine entrenchment continua along the two dimensions that structure the Euclidean space.

G. Desagulier

## 5. CONCLUSION

- Need for better statistics in the study of the collocation patterns of moderator  
-> collocation analysis;
- Combining univariate and multivariate statistics helps map usage patterns and conceptual structure  
-> the relationship between moderators and adjectives is indeed bidirectional and can be represented graphically;
- Moderators are indeed similar and different BUT  
-> moderators have a functional basis in common but do not always operate within the same conceptual domains. If they do, they follow a division of labor;
- Instead of clustering lexical co-occurrence data alone, it will be in the interest of future research to integrate more strata of richly annotated data within the same plot (e.g. information concerning grading force and semantic classes of adjectives). The resulting two-dimensional map will provide a finer-grained representation of the scale of synonymy of moderators.

## 6. BIBLIOGRAPHY

- 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: John Benjamins.
- 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(4), 554-565.
- Benzécri, J. (1973). *L'analyse des données, 2. L'analyse des correspondances*. Paris Bruxelles Montréal: Dunod.
- Benzécri, J. (1984). *Analyse des correspondances, exposé élémentaire* (2<sup>nd</sup> ed.). Paris: Dunod.
- Cruse, D. A. (1986). *Lexical semantics*. Cambridge Cambridgeshire; New York: Cambridge University Press.
- Cruse, D. A. (1995). *Lexical semantics*. Cambridge Cambridgeshire; New York: Cambridge University Press.
- Dagmar, D. (2006). Ways of intending: Delineating and structuring near-synonyms. *Corpora in cognitive linguistics* (pp. 19-56) Mouton de Gruyter.
- Davies, M. (1990-present). *The Corpus of Contemporary American English (COCA): 410+ million words*. <http://corpus.byu.edu/coca>
- Divjak, D. (2010). *Structuring the Lexicon: A Clustered Model for Near-Synonymy*. Berlin; New York: De Gruyter Mouton.
- Divjak, D., & Gries, S. T. (2008). Clusters in the mind? Converging evidence from near synonymy in Russian. *The Mental Lexicon*, 3(2), 188-213.
- Downing, A., & P. Locke (2006). *English grammar: A university course* (2<sup>nd</sup> ed.). London: Routledge.
- 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*, 1-3.
- 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(5), 219-224.
- Goldberg, A. E. (2006). *Constructions at work: the nature of generalization in language*. Oxford; New York: Oxford University Press.

- Greenacre, M. J. (2007). *Correspondence Analysis in Practice* (2<sup>nd</sup> ed.). Boca Raton: Chapman & Hall/CRC.
- Gries, S. T. (2007). *Coll.analysis 3.2. A program for R for windows 2.x*.
- Gries, S. T., & Stefanowitsch, A. (2004). Extending collocation analysis: A corpus-based perspective on 'alternations'. *International Journal of Corpus Linguistics*, 9(1), 97-129.
- Gries, S. T., & 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*. Stanford, CA: CSLI.
- Gries, S. T., & Stefanowitsch, A. (2006). *Corpora in Cognitive Linguistics: Corpus-Based Approaches to Syntax and Lexis*. Berlin; New York: Mouton de Gruyter.
- Kennedy, G. (2003). Amplifier Collocations in the *British National Corpus*: Implications For English Language Teaching. *TESOL Quarterly*, 37(3), 467-487.
- Kilgarriff, A. (2001). Comparing corpora. *International Journal of Corpus Linguistics*, 6(1), 97-133.
- Langacker, R. W. (1987). *Foundations of Cognitive Grammar*. Stanford, California: Stanford University Press.
- Langacker, R. W. (1991). *Foundations of Cognitive Grammar*. Stanford, California; Stanford University Press.
- Langacker, R. W. (2008). *Cognitive Grammar: a Basic Introduction*. Oxford; New York: Oxford University Press.
- 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.
- Nevalainen, T., & Rissanen, M. (2002). *Fairly pretty or pretty fair?* On the development and grammaticalization of English downtoners. *Language Sciences*, 24(3-4), 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.
- Quine, W. V. (1951). Main trends in recent philosophy: Two dogmas of empiricism. *The Philosophical Review*, 60(1), pp. 20-43.
- Simon-Vandenberg, A. (2008). *Almost certainly and most definitely*: Degree modifiers and epistemic stance. *Journal of Pragmatics*, 40(9), 1521-1542.
- Stefanowitsch, A., & Gries, S. T. (2003). Collocations: investigating the interaction of words and constructions. *International Journal of Corpus Linguistics*, 8(2), 209-243.
- Stoffel, C. (1901). *Intensives and downtoners: A study in English adverbs*. Heidelberg: Carl Winter.

---

<sup>i</sup> Cruse's definition of cognitive synonymy echoes Quine's (1951). In later editions of Cruse's monograph, cognitive synonymy is relabeled 'propositional synonymy' (Cruse, 1995).