Analyzing Diachronic Change in the American English Amplifier System

Dr. Martin Schweinberger
www.martinschweinberger.de

ICAME 39: Tampere

Habilitation (in progress)
Acquisition, Variation, and Diachronic Development of Intensification in English

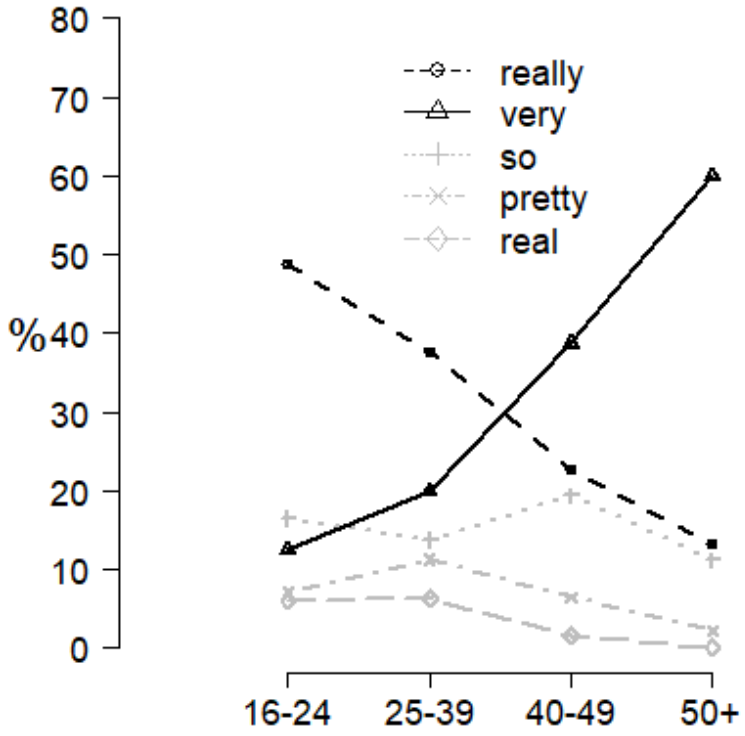
Intensification is related to the semantic category of *degree* (degree adverbs) and ranges between very low intensity (downtoning) and very high (amplifiers) (Quirk et al. 1985: 589–590).

- Amplifiers (Tagliamonte 2008)
  - Maximizers (e.g. *completely*)
  - Boosters (e.g. *very much*)
- Downtoners
  - Approximators (e.g. *almost*)
  - Compromisers (e.g. *more or less*)
  - Diminishers (e.g. *partly*)
  - Minimizers (e.g. *hardly*)

Previous Research

- Intensification is interesting to analyze because. . .
  - major area of grammatical change (cf. Brinton and Arnovick 2006: 441)
  - crucial for the "social and emotional expression of speakers" (Ito and Tagliamonte 2003: 258)
  - amplifying *really* replaces *very* (lexical replacement)
    - NZE (cf. D'Arcy 2015)
    - BrE (cf. Ito and Tagliamonte 2003)
    - CanE (cf. Tagliamonte 2005, 2008)
    - IrE (cf. Schweinberger sub)

Previous study of intensification in New Zealand English (Schweinberger sub)



Amplification in ICE NZ

Q_1:

Do the COHA data mirror the trend that *really* is replacing *very*?

Data Processing

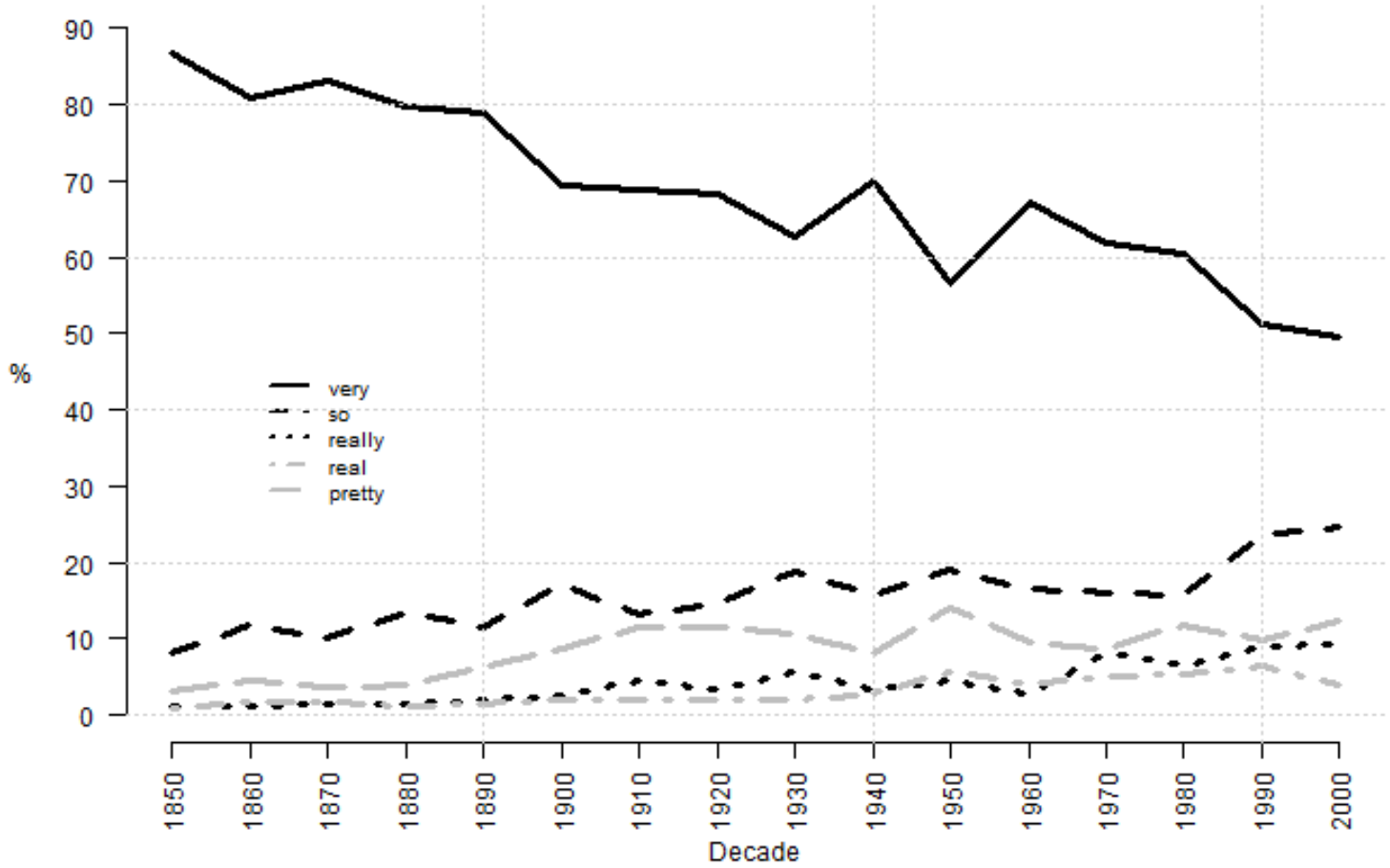
- Extraction of adjectives based on COHA POS tags
- Determine which adjectives were intensified (amplifier occurred before adjective)
- Removal of adjectives that did not occur before a full stop or before a noun (to determine syn. function; predicative vs attributive)
- Removal of adjectives that were not intensified by at least two different amplifier types

Data Processing

- Removal of adjectives that were intensified in less than five percent of cases
- Removal of negated adjectives
- Removal of comparative and superlative forms
- Only fiction data!

Data Summary: Intensifiers COHA data

Intensifier	**N**	**%**	**Int.** (%)
∅	148,560	92.08	
very	6,819	4.23	53.36
so	1,538	0.95	12.03
pretty	848	0.53	6.64
really	407	0.25	3.18
real	288	0.18	2.25
perfectly	249	0.15	1.95
extremely	222	0.14	1.74
mighty	215	0.13	1.68
entirely	212	0.13	1.66
highly	189	0.12	1.48
awfully	127	0.08	0.99
particularly	120	0.07	0.94
terribly	106	0.07	0.83
truly, completely, exceedingly, awful, remarkably, totally, wholly, utterly, dead, especially, absolutely, distinctly, incredibly, decidedly, intensely, extraordinarily, purely, dreadfully, obviously, wildly, positively, exceptionally, genuinely, enormously, immensely, profoundly, seriously, strikingly, amazingly, infinitely, clearly, plenty, fucking, true, well, actually, bloody, certainly, frightfully, hopelessly	100-10 (1.343)	0.83	10.5
specially, heavily, strongly, considerably, fully, hugely, crazy, definitely, greatly, shocking, badly, complete, fiercely, grossly, immediately, openly, surely, excruciatingly, overwhelmingly, overwhelmingly, super, terrifically, mad, wicked	10- (97)	0.01	0.77
Total	161,340	7.92	100



Amplifier Types against Time (COHA)

Q_2:

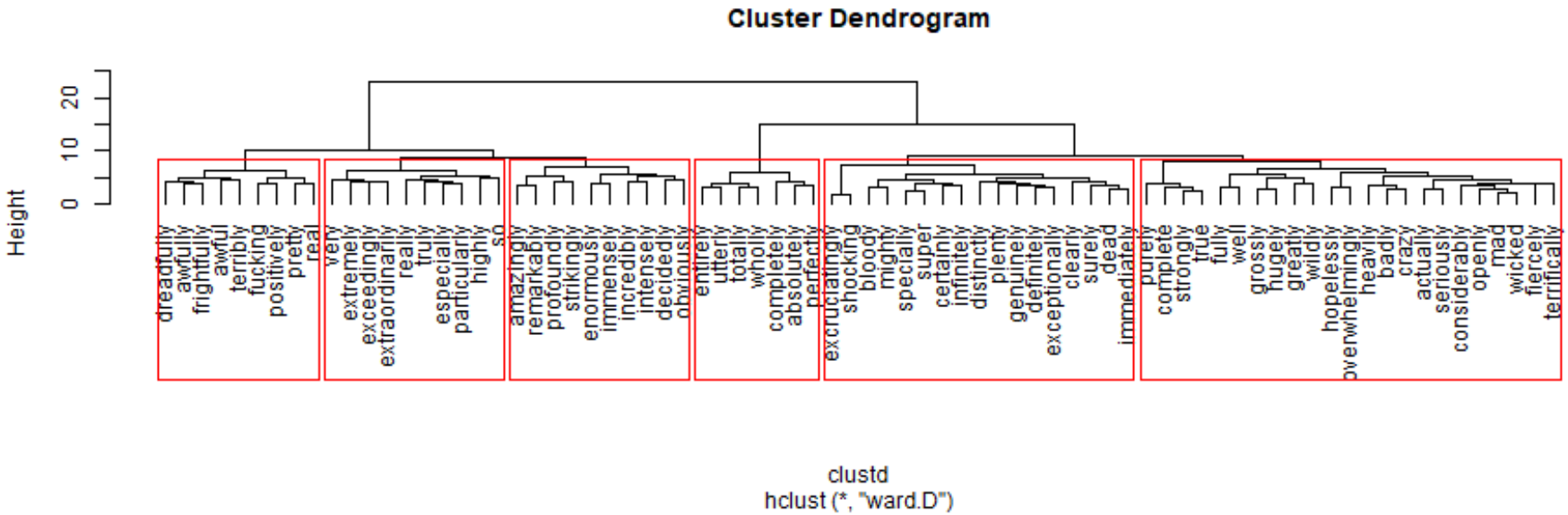
Why is *so* (and not *really*) replacing *very* in the COHA data?

H_1:

Semantic similarity

so is replacing *very* in AmE because they have similar collocational profiles
(*really* fails at replacing *very* because their collocational profiles are notably different).

Semantic similarity



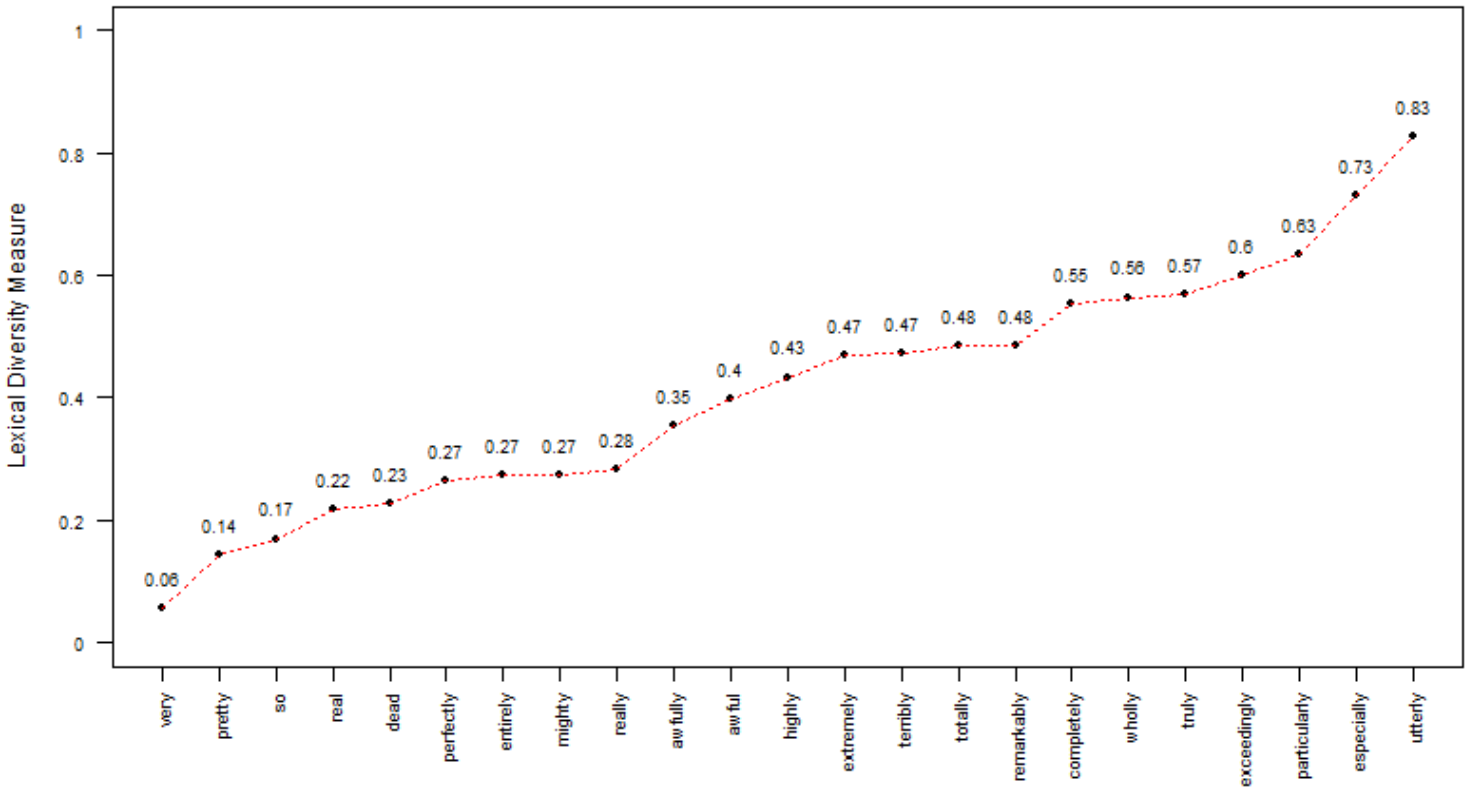
Dendrogram of Semantic Vector Space Model (Levshina 2015) based on Distribution of Amplifiers and Adjectives

H_2:

Semantically bleached

so is replacing *very* in AmE because it
is more bleached (less collocationally restricted) than
really while being still more expressive than *very*.

Semantically bleached



Semantic Diversity: Amplifier Tokens against Adjective Types ($N_{Adj.Types}/N_{Amp.}$)

H_3:

Collocation with High Frequency Adjectives

so is replacing *very* in AmE because it attaches to high frequency adjectives. Thereby *so* is becoming more deeply entrenched and thus easier retrieved from the lexicon.

Adjective	Intensifier	1850	1880	1910	1940	1970	2000
different	other	9.7	21.1	57.9	45.5	48.1	64.9
good	other	8.2	17.5	19.4	11.7	11.2	13.8
pretty	other	5	13.6	25	13	25	33.3
good	pretty	11.5	14	30.6	22.5	26.4	28.5
happy	pretty	0	0	0	0	0	12.5
nice	pretty	0	4.3	0	8.6	0	2.9
bad	real	0	0	0	0	16.7	20
good	real	3.3	2.6	2.8	9.2	8	4.6
happy	real	0	0	0	0	0	12.5
nice	real	7.7	4.3	0	11.4	9.4	5.7
pretty	real	0	0	0	0	12.5	8.3
bad	really	0	2.8	0	0	8.3	17.8
good	really	4.9	3.5	6.5	2.5	7.2	10.8
important	really	0	0	0	18.2	12.5	12.5
nice	really	0	0	5.6	2.9	6.2	17.1
beautiful	so	18.2	11.1	14.3	31.8	15.4	33.3
different	so	3.2	7.9	10.5	9.1	7.4	2.7
good	so	0	6.1	5.6	5.8	4.8	10
happy	so	27.3	31.8	31.2	30.8	10	37.5
pretty	so	0	9.1	12.5	4.3	25	25
bad	very	76.5	69.4	50	50	54.2	15.6
beautiful	very	63.6	66.7	78.6	45.5	50	40.7
different	very	87.1	71.1	31.6	45.5	44.4	32.4
good	very	72.1	56.1	35.2	48.3	42.4	32.3
happy	very	63.6	54.5	50	38.5	80	25
important	very	80	100	59.1	63.6	87.5	56.2
nice	very	76.9	73.9	61.1	57.1	65.6	54.3
pretty	very	95	77.3	62.5	82.6	37.5	25

decreasing trend (linear model, $p<.05$)
increasing trend (linear model, $p<.05$)

Loss of collocates

- bad: very → real
- beautiful: very → (so)
- different: very → other
- nice: very → really
- pretty: very → so
- good: very → (so) !

	Adjectives						
	good	beautiful	different	bad	happy	quiet	Sum All
Abs. Freq.	31,184	6,199	5,883	5,544	4,406	3,350	161,340

Summary, Problems & Outlook

Discussion

Main points

- *so* is replacing *very* in the COHA data because . . .
  - it is semantically similar to *very* (like *really* but unlike *pretty*)
  - it is more bleached (than *really* but less so than *pretty*) while being still more expressive than *very*.
  - it attaches to high frequency adjectives and is thereby becoming more deeply entrenched and thus easier retrieved from the lexicon.
- In conclusion, the state of rivalry in the AmE amplifier system is ongoing and none of the variants is dominant yet but *so* is the most likely candidate for becoming dominant.

Discussion

Problems

- data sets does not yet contain registers other than fiction (genre comparisons not yet possible !)
- non-intensifiers are still present in the data (e.g. *quite*)
- no differentiation between boosters and maximizers
- no manual cross–evaluation of automated classification

Outlook

- further investigation of collocational patterns (across time) using vector space models
- Regression modeling of the rise of *so* (done but no time)

Thank you so, really, very much!

Brinton, L. J. and L. K. Arnovick (2006). *The English Language: A Linguistic History*. Oxford: Oxford University Press.

D'Arcy, A. F. (2015). Stability, stasis and change - the longue duree of intensification. *Diachronica 32*(4), 449–493.

Ito, R. and S. Tagliamonte (2003). Well weird, right dodgy, very strange, really cool: Layering and recycling in english intensifiers. *Language in Society 32*, 257–279.

Levshina, N. (2015). *How to do linguistics with R: Data exploration and statistical analysis*. Amsterdam: John Benjamins Publishing Company.

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

Schweinberger, M. (sub.). Using semantic vector space models to investigate lexical replacement – a corpus based study of ongoing changes in intensifier systems. In Y. Asahi (Ed.), *Proceedings of Methods in Dialectology XVI*, pp. NA. NA: NA.

Tagliamonte, S. (2005). So who? like how? just what?: Discourse markers in the conversations of young canadians. *Journal of Pragmatics 37*(11), 1896–1915.

Tagliamonte, S. (2008). So different and pretty cool! recycling intensifiers in toronto, canada. *English Language and Linguistics 12*(2), 361–394.

Analyzing Diachronic Change in the American English Amplifier System

Dr. Martin Schweinberger
www.martinschweinberger.de

ICAME 39: Tampere

Appendix

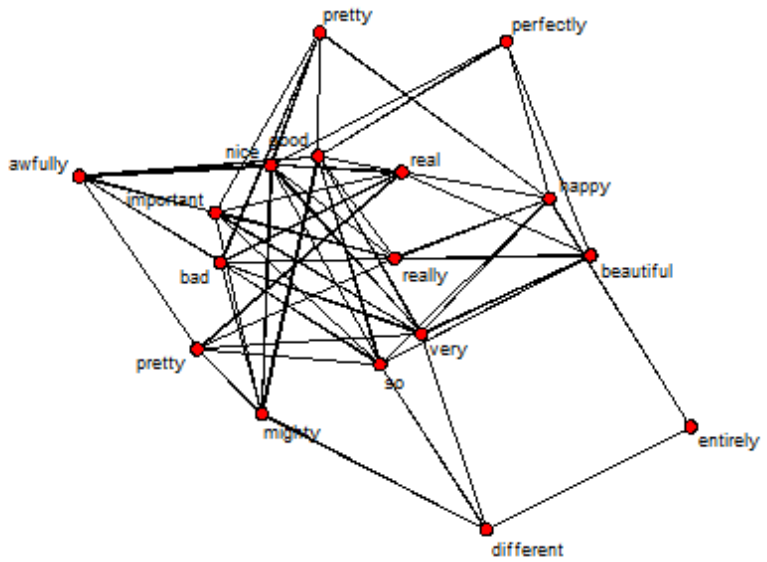
Collocations by Decade

Results: Distinctive Collexeme Analyses

Age	Adjective	Intensifier	OddsRatio	Bonf. Corr. Sig
1880	pretty	good	11.25	p<.001***
1900	pretty	bad	9.96	p<.01**
1900	pretty	good	6.45	p<.001***
1900	so	glad	32.44	p<.001***
1910	pretty	good	8.87	p<.001***
1920	pretty	good	9.80	p<.001***
1930	pretty	good	4.00	p<.001***
1940	pretty	good	7.84	p<.001***
1950	pretty	good	9.08	p<.001***
1960	pretty	good	10.38	p<.001***
1970	pretty	good	11.39	p<.001***
1980	pretty	good	6.25	p<.001***
1990	pretty	good	7.85	p<.001***
1990	so	sorry	8.07	p<.01**
2000	pretty	good	6.69	p<.001***
2000	real	bad	12.46	p<.01**
2000	so	sorry	9.48	p<.001***

p-values obtained by Fisher's Exact tests

Semantically bleached



Network Analysis plot of Amplifier and Adjective Types.