Analyzing historical changes in Irish English adjective amplification

slides available at

www.martinschweinberger.de

m.schweinberger@uq.edu.au

R code upon request

Dr. Martin Schweinberger

The University
Of Queensland
AUSTRALIA

Phenomenon: Adjective Amplification

(1) Adjective amplification in predicative contexts

a. they were very kind to me on bord (HCIE:Alderd01:1880)

b. John was prety sick all spring (HCIE:CarroA36:1887)

c. I am so glad (HCIE:Burke:03:1884)

d. I was really glad to hear that (HCIE:Spratt01:1869)

(2) Adjective amplification in attributive contexts

a. I have very good health (HCIE:GambJ:06:1876)

b. I have pretty good wages at present (HCIE:ButleJ02:1850)

c. we have so large grounds to pley (HCIE:Sloane04:1892)

d. I got a real nice Passage (HCIE:Markey01:1889)

Phenomenon: Adjective Amplification

- Subtype of intensification
- Related to the semantic category of *degree* (degree adverbs) and ranges from low (downtoning) to high (amplifiers) (Quirk et al. 1985: 589–590)
  - Amplifiers
    - Boosters, e.g. *very*
    - Maximizers, e.g. *completely*
  - Downtoners
    - Approximators, e.g. *almost*
    - Compromisers, e.g. *more or less*
    - Diminishers, e.g. *partly*
    - Minimizers, e.g. *hardly*

Previous Research

Amplification

- Extensive history of research and a substantial amount of corpus-based research on intensification (e.g Aijmer 2011, 2018; Bauer and Bauer 2002; Borst 1902; Bolinger 1972; D'Arcy 2015; Fuchs 2016, 2017; Ito and Tagliamonte 2003; Núñez Pertejo and Palacios 2014; Palacios and Núñez Pertejo 2012; Tagliamonte and Roberts 2005; Macaulay 2006; Tagliamonte 2006, 2008)
- Intensification is considered a major area of grammatical change in English which makes it an ideal test case for processes of language change (cf. Brinton and Arnovick 2006: 441)
- Crucial for "social and emotional expression of speakers" (Ito and Tagliamonte 2003: 258)

Previous Research

- Amplifying *really* replaces *very* (lexical replacement)

  (see D'Arcy (2015) for NZE; see Ito and Tagliamonte (2003) and Barnfield and Buchstaller (2010) for North East British English, Tagliamonte (2008) and Tagliamonte and Denis (2014) for Toronto English; see Tagliamonte and Denis (2014) for South Eastern Ontario English)

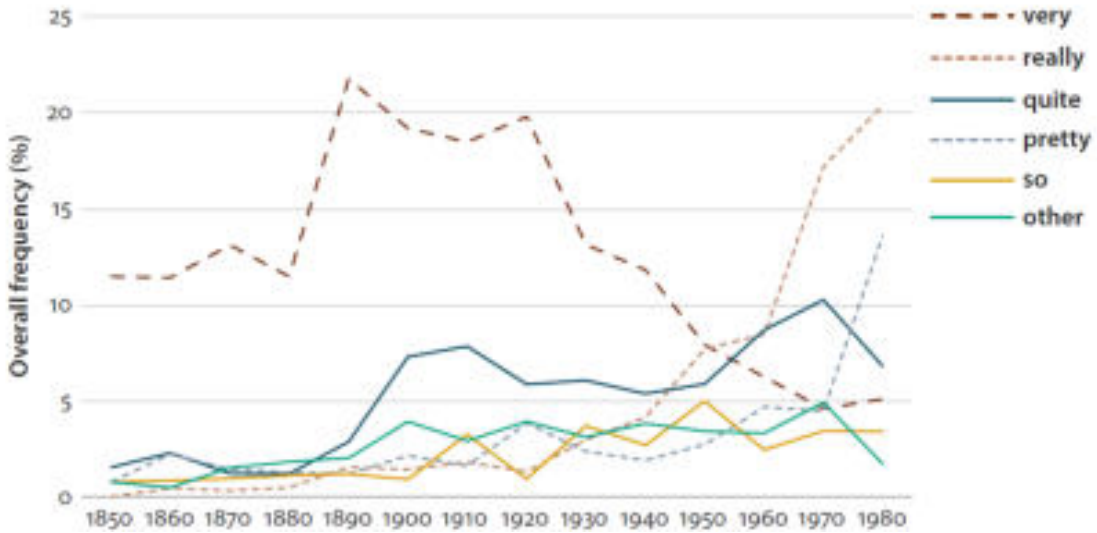


Figure 1: Adapted from D'Arcy (2015: 468)

Previous Research

- Amplifying *really* replaces *very* (lexical replacement)

(see Schweinberger (subm) for Irish English)

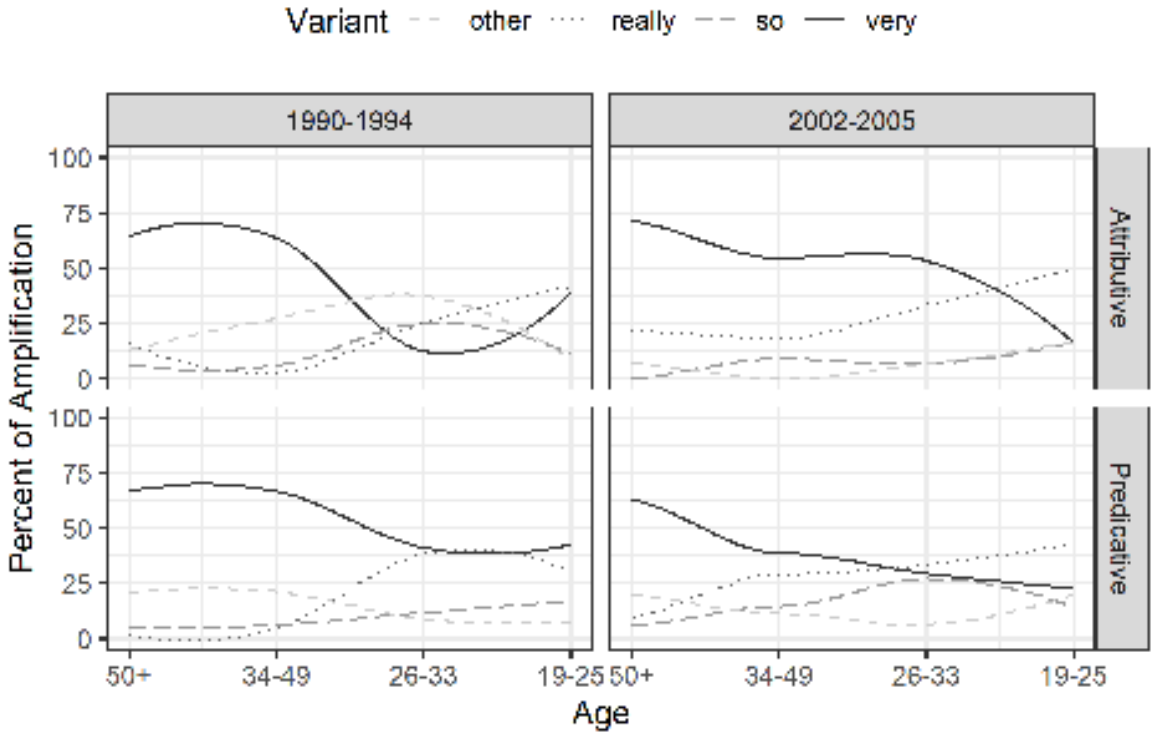


Figure 2: Adapted from Schweinberger (subm)

Motivation

Adjective amplification has been described as a

- "major area of grammatical change" (cf. Brinton and Arnovick 2006: 441)
- "[...] waxing and waning of forms, alongside invention and renewal" (D'Arcy 2015: 450)
- "Competition, change and recycling among intensifiers have been going on in English since the Old English period" (Tagliamonte 2008: 361)
- "in this sea of change, processes of expansion and contraction are occurring all the time" (Partington 1993: 180)

Research Question

Q

How did the adjective amplifier system in Irish English evolve during the last 300+ years?

→ Is adjective amplification really a site of perpetual waxing and waning of forms?

Data and Methodology

the Hamburg Corpus of Irish English (HCIE)

Hamburg Corpus of Irish English (HCIE)

- Letters written by Irish emigrants between 1675 and 1935
- App. 1,000 letters amounting to app. 500,000 words
- Accompanied by a limited range of metadata (e.g. year of composition) and information about the "speakers" (gender)

Period	**Letters (N)**	**Words (N)**
1675-1750	16	9,818
1751-1850	355	225,032
1901-1935	608	322,107
Unknown	1	223
Total	**980**	**557,180**

Hamburg Corpus of Irish English (HCIE)

- Semi-literate authors (dictated or self-written) who used a profoundly speech-like style (the most speech-like data we will get. . .)
- Lower social strata (not upper-class, highly literate and multilingual writers as in the INNSBRUC letter corpus)
- Letters predominantly from the US, but also from Australia and New Zealand (but substantially fewer)

Data Processing

Data Processing

- Cleaned and automatically corrected/spell-checked the letters
- Part–of-speech tagged (OpenNLP vis `R`) the cleaned letters
- Retrieved adjectives (PoS–tag JJ)
- Determined whether adjective were preceded by an amplifier (member of a predefined set of amplifiers)
- Manual cross–evaluation of automated classification
- Metadata and speaker information

Data Processing

Removed...

- negated adjectives
- comparative and superlative forms
- adjectives that were not amplified by at least two different amplifier types
- adjectives that were preceded by downtoners
- strange forms (e.g. *much*)

Statistical procedures

Lexical diversity scores

Dividing the number of adjective types a given amplifier co-occurs with by the number of tokens of that amplifier type (see 3).

(3) $\text{LD} = \text{N}_{Adj.Types}/\text{N}_{Amp.Tokens}$

LD can reach a maximum value of 1 (high lexical diversity). The lower the LD value, the lower the degree of lexical diversity.

Amplifier	Amp. Tokens (N)	Adj. Types (N)	Calculation	LD
variant_A	10	1	1/10	.1
variant_B	10	5	5/10	.5
variant_C	10	10	10/10	1
$\text{very}_{1675-1750}$	11	17	11/18	.647
$\text{very}_{1751-1850}$	86	368	86/368	.234
$\text{very}_{1851-1930}$	102	845	102/845	.121

Covarying collexeme analysis

- Member of the collostructional family of analyses (Hilpert (2006); Stefanowitsch and Gries (2005))
- Quantifies and evaluates of attraction between elements that occur in two distinct slots within a specified construction (here: 1^{st} slot = amplifier slot; 2^{nd} slot = adjective slot)
- How much more likely is it that *nice* occurs in the 2^{nd} slot given that *really* occurs in the 1^{st} slot compared with another amplifier taking the 1^{st} slot?
- High values = strong attraction: the amplifier and adjective co-occur more often together than would be expected by chance;
  Low values = the amplifier and adjective co-occur less often together than would be expected by chance.

Semantic vector space models (Levshina 2015)

Semantic Vector Space Models (SVSM) are based on co-occurrence profiles of adjectives and amplifiers

- Idea: words that have similar collocational profiles are similar in meaning; e.g. amplifiers that co-occur with the same adjectives are semantically similar
- Advantages of such models are that SVSM ...
  - provide a data-driven measure of semantic similarity (do not rely on any theoretical framework)
  - take variety specific differences in use/meaning into account (fine-grained classifications)
  - inform about statistical significance)

Semantic vector space models

Example

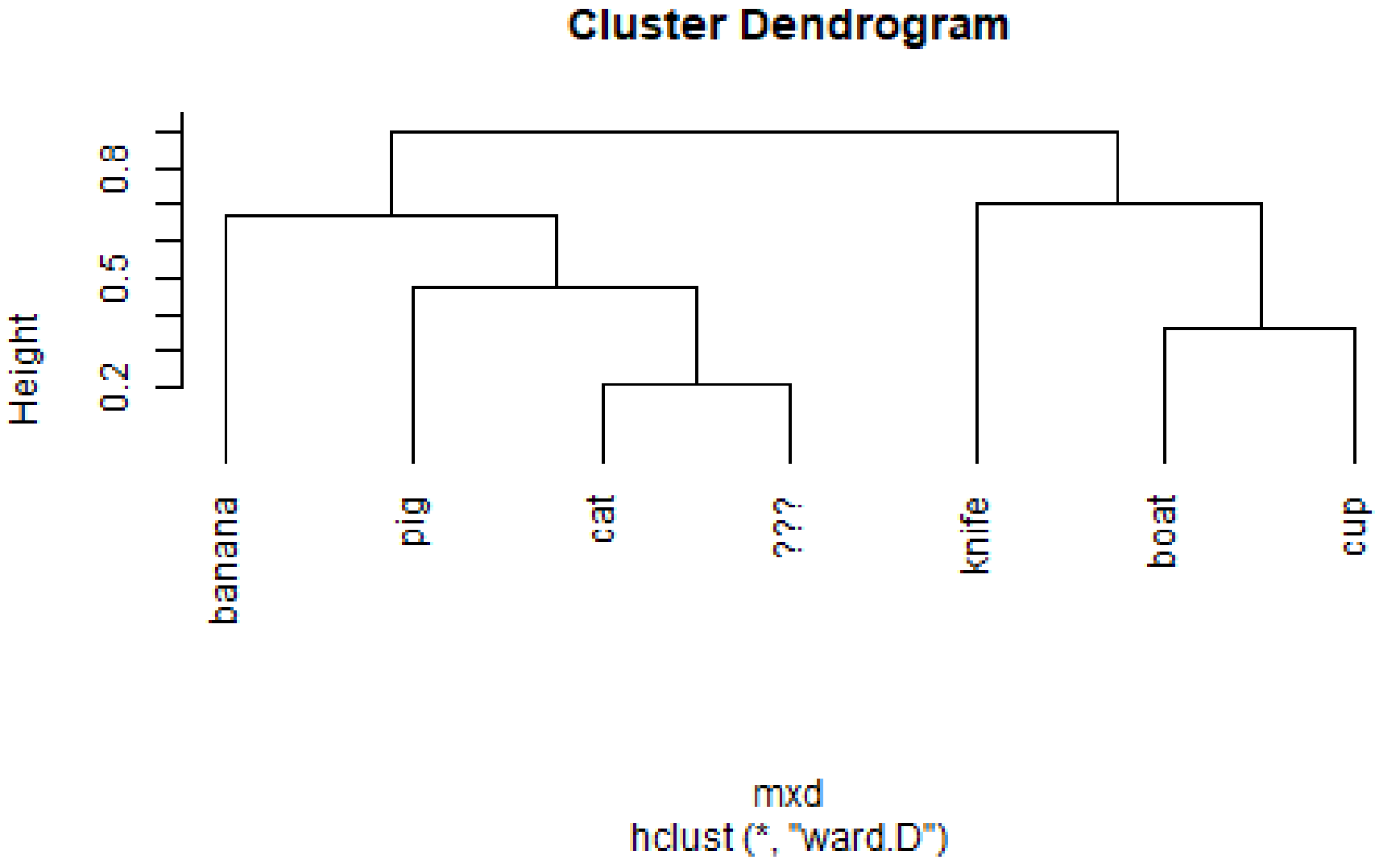
	get	**see**	**use**	**hear**	**eat**	**kill**
knife	31	16	69	0	2	0
cat	36	38	4	4	6	20
???	66	58	9	34	28	12
boat	46	21	17	4	0	0
cup	59	6	5	1	1	0
pig	4	15	3	1	7	21
banana	7	2	2	0	12	0

Table 1: Co-occurrences of selected nouns and verbs in the British National Corpus.

Semantic vector space models

	knife	**cat**	**???**	**boat**	**cup**	**pig**
cat	.62					
???	.60	.21				
boat	.48	.33	.32			
cup	.76	.59	.58	.36		
pig	.72	.36	.45	.64	.88	
banana	.71	.57	.47	.60	.72	.64

Table 2: Distance matrix based on co-occurrences of selected nouns and verbs in the British National Corpus.


Cluster Dendrogram
Height
0.2
0.5
0.8
banana
pig
cat
???
knife
boat
cup
mxd
hclust (*, "ward.D")

Results

Variants across real-time

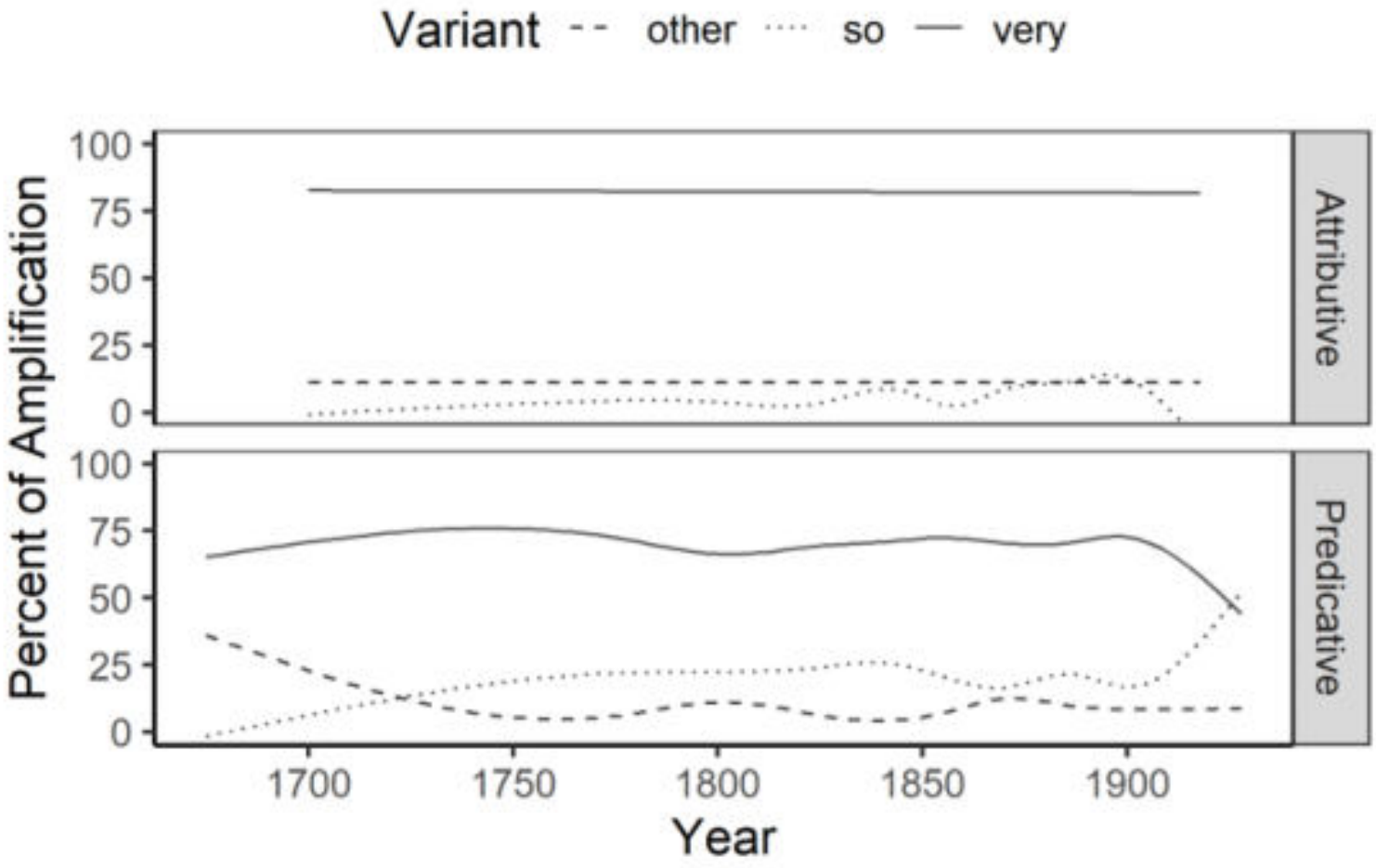


Figure 3: Percent of amplifier variants in IrE (HCIE) across real-time.

Interpretation

Variants across real-time

- Attributive contexts
  - Remarkable(!) stability and stasis
- Predicative contexts
  - Little change in until 1900
  - 1900 and later: *so* replaces *very*

Q

Is the replacement of *very* by *so*
accompanied by changes in co-occurrence patterns?

Scenario 1 (Broadening)

so increases because it associates with many (but infrequent) adj. types

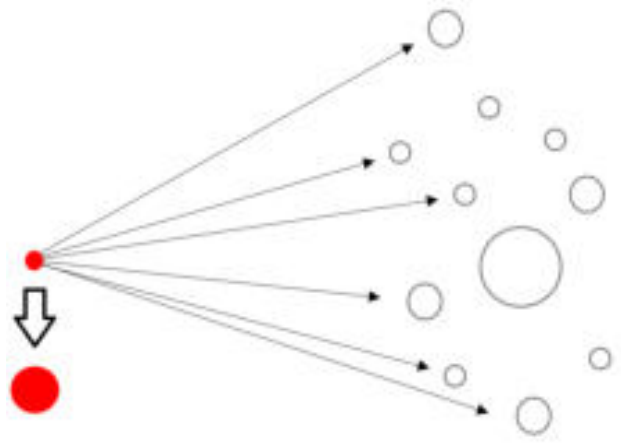
(Mair 2004: "delayed increase of discourse frequency" hypothesis)

Prediction

Co-occurrence with many different adjective types

→ high lexical diversity

→ weak coll. attraction with specific adj. types



Scenario 2 (Specialization)

so increases because it associates with few but frequent adj. types (HFAs)

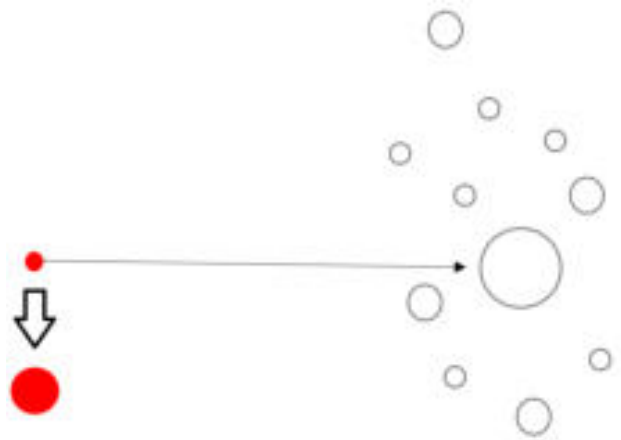
(Lorenz 2002: 144; Méndez-Naya 2003: 375; Tagliamonte and Roberts 2005: 285)

Prediction

Co-occurrence with few high frequency adjectives

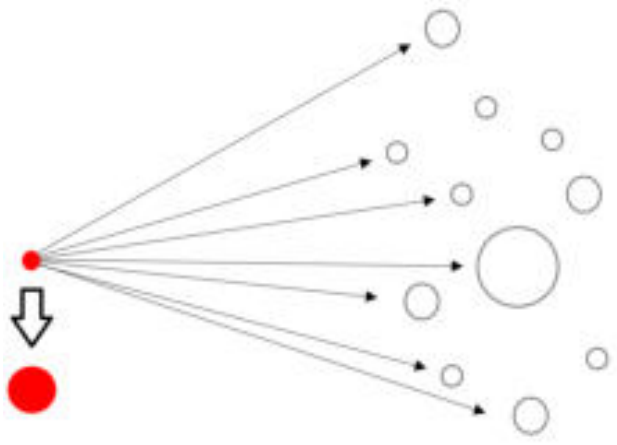
→ low lexical diversity

→ strong coll. attraction with high-freq. adj. types



Scenario 3 (Randomness)

so increases because it associates with random adj. types
→ We cannot predict which variants become successful based on their coll. profile.



Lexical diversity scores of amplifier variants across real-time

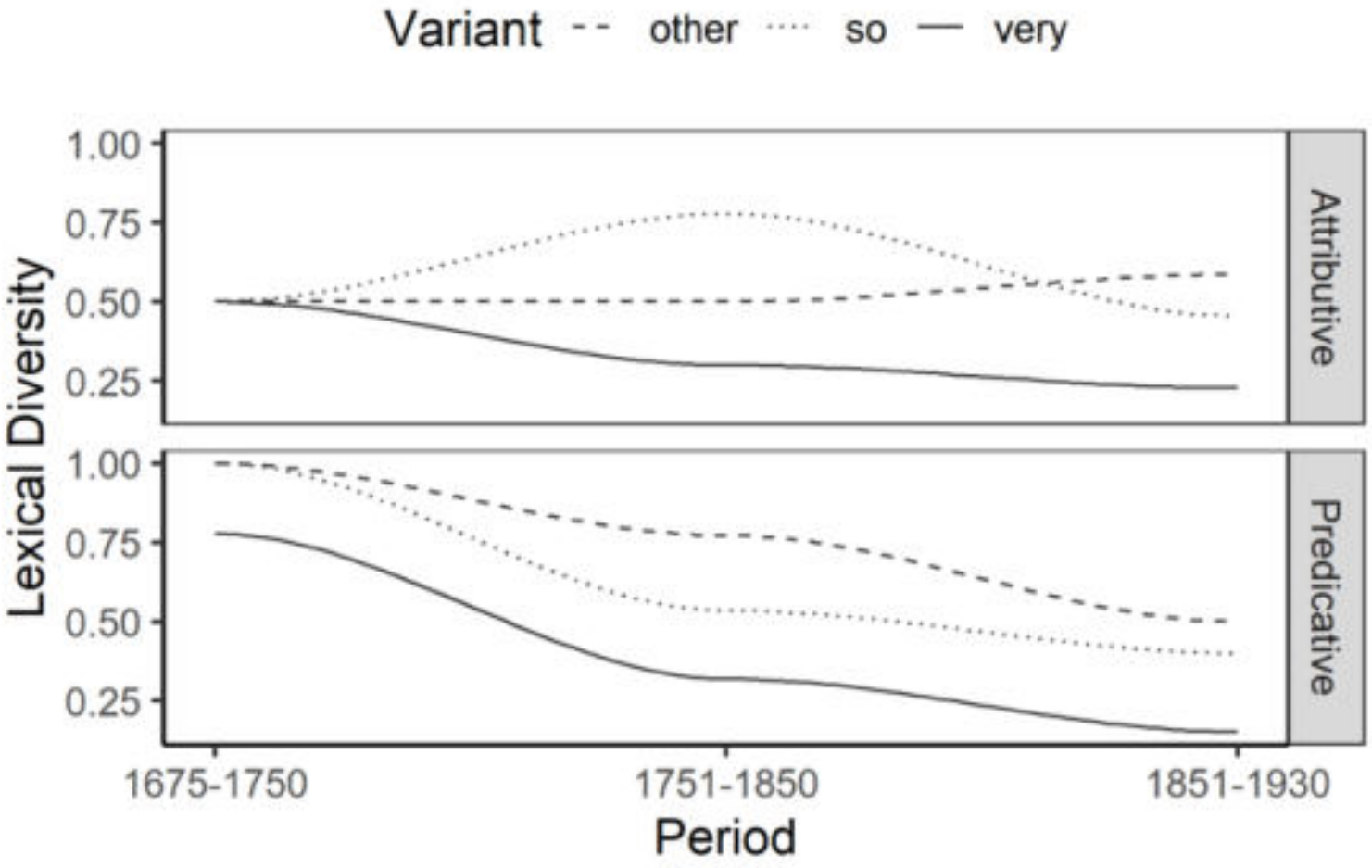


Figure 4: Lexical diversity scores in IrE (HCIE) across real-time.

Interpretation

The increase in *so* is not accompanied by any discernible increase or decrease in lexical diversity...

Q

Could the change in the amplifier system
be triggered by changes in the frequency of adjectives?

Change in adjective types across real-time

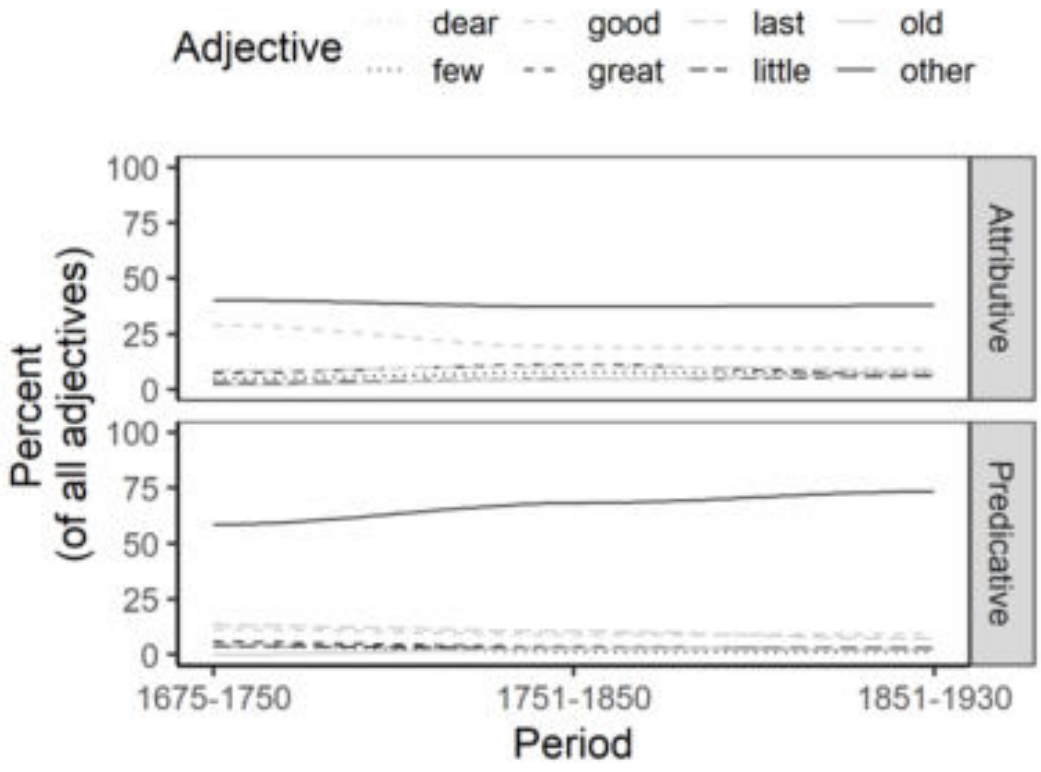


Figure 5: Adjective types in IrE (HCIE) across real-time.

Interpretation

The increase in *so* is not triggered by any discernible change in the frequencies of adjectives. . .

Q

Could the change in the amplifier system be based on changes in the collocation profile of amplifier-adjective bigrams?

Changes in collocation strength across real-time (covarying collexeme analysis)

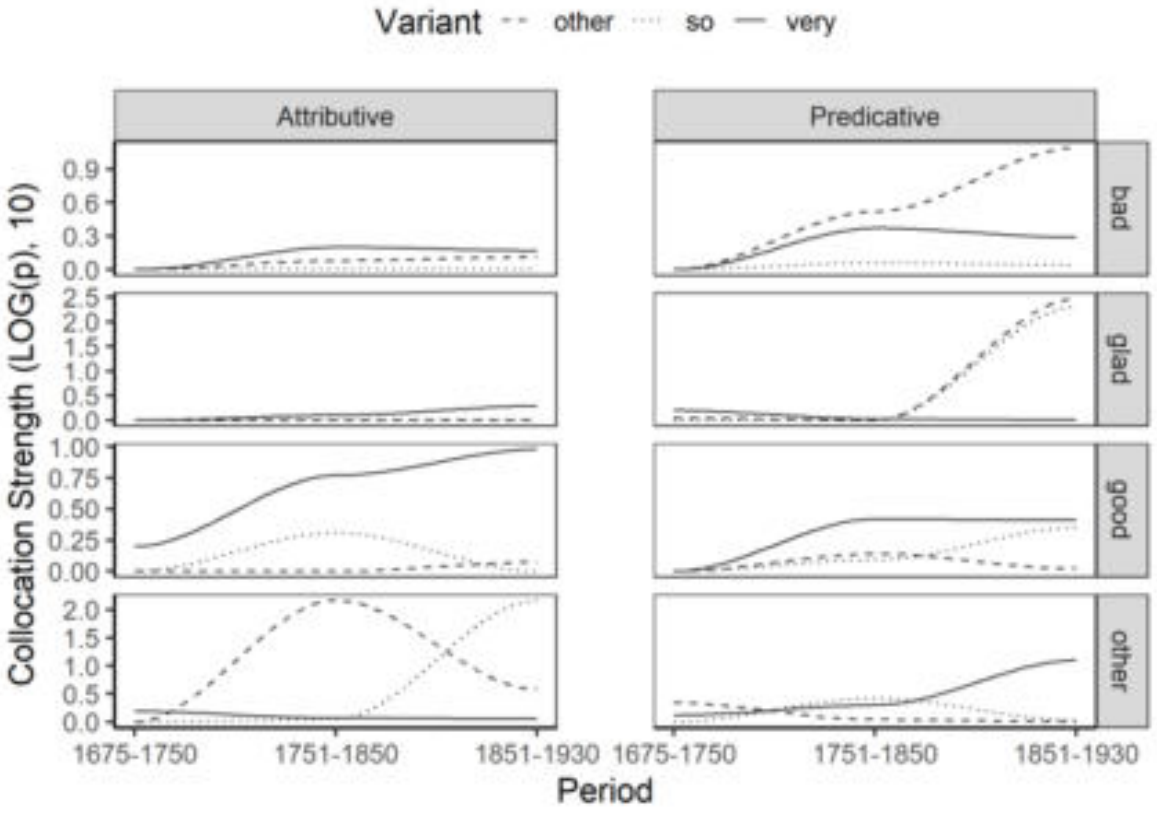


Figure 6: Collocation strength in IrE (HCIE) across real-time.

Interpretation

In the newest data, *so* is increasingly co-occurring with *good* (the most frequent adjective) and *glad* (another HFA). These changes in collocation strength are, however, relatively small and not significant...

Q

Could the change in the amplifier system be dependent upon semantic similarity?

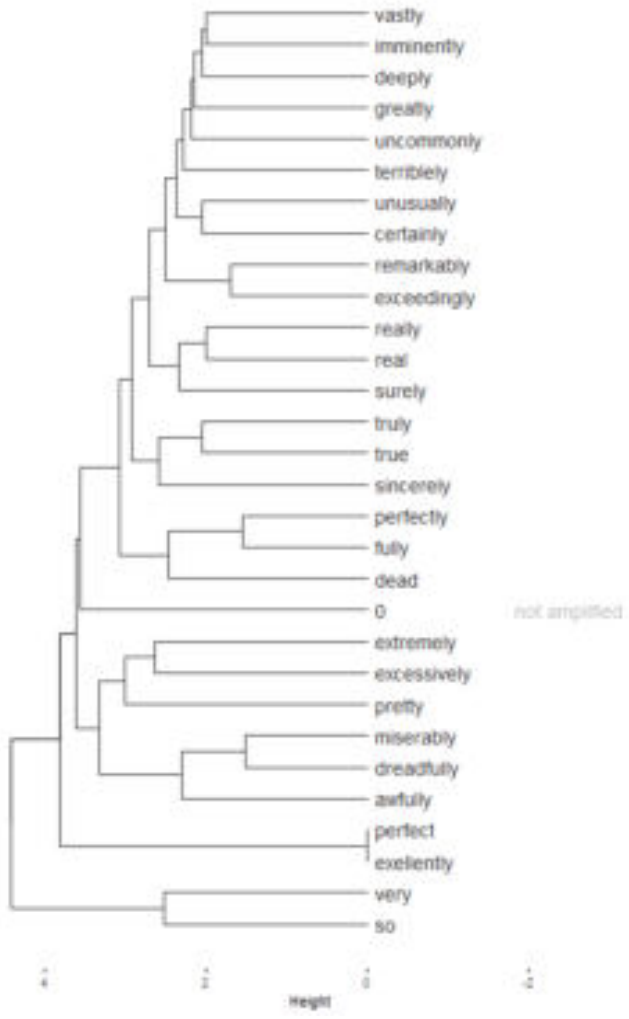


Figure 7: Semantic similarity of amplifier types in IrE (HCIE).

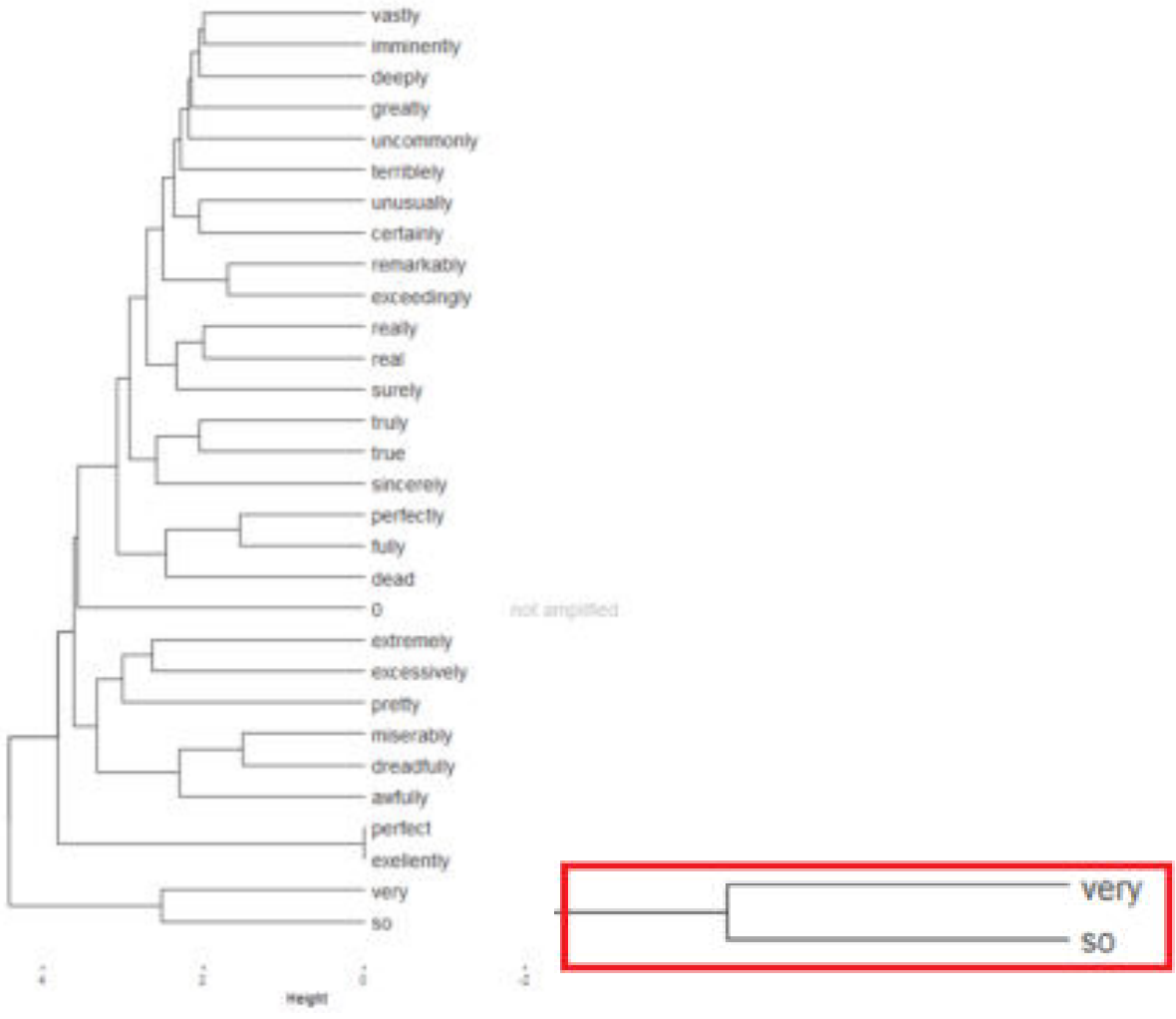


Figure 8: Semantic similarity of amplifier types in IrE (HCIE).

Interpretation

so and *very* are, indeed, most similar with respect to their collocational profiles. This suggest that replacement is dependent upon semantic similarity (Schweinberger (fcb))! However, this does not explain why so replaces very in the HCIE data in the first place...

Q

Is this pattern unique to Irish English?

→ Is this an isolated phenomenon?

Schweinberger (fca)

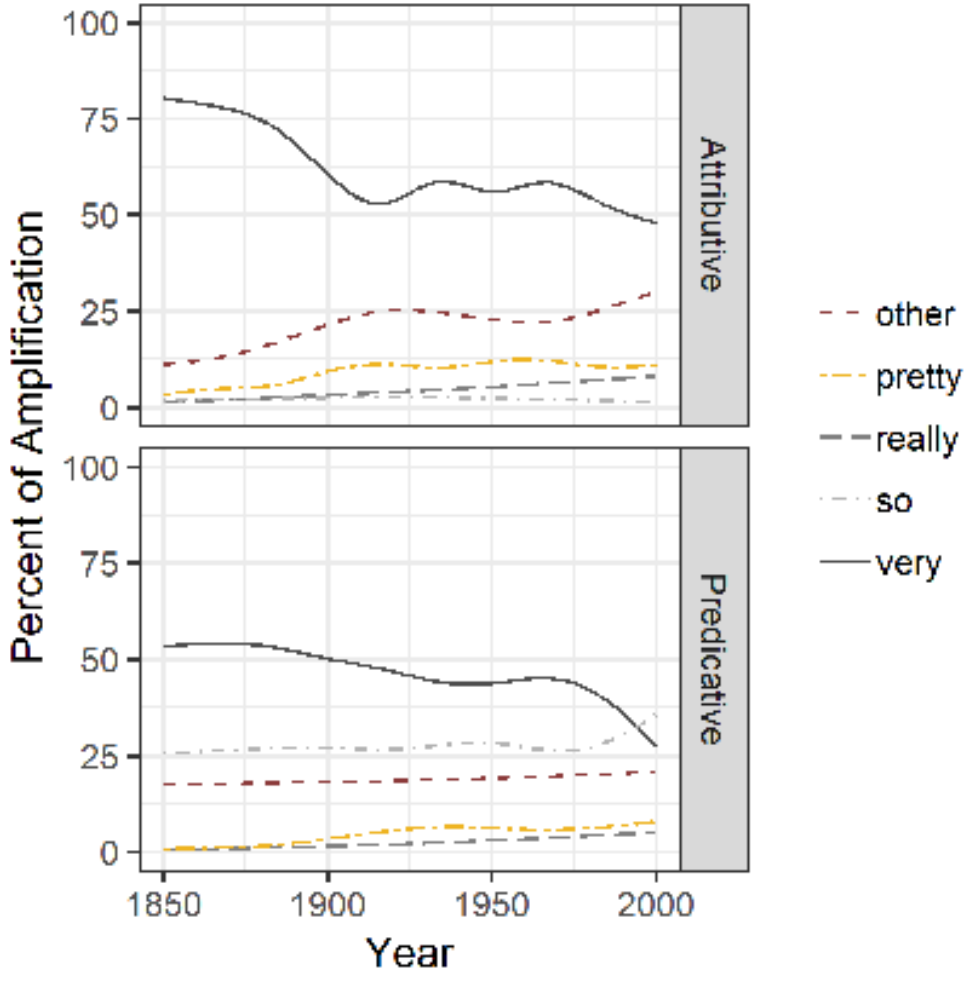


Figure 9: Percent of amplifier variants in AmE (COHA) across real-time. Schweinberger (fca).

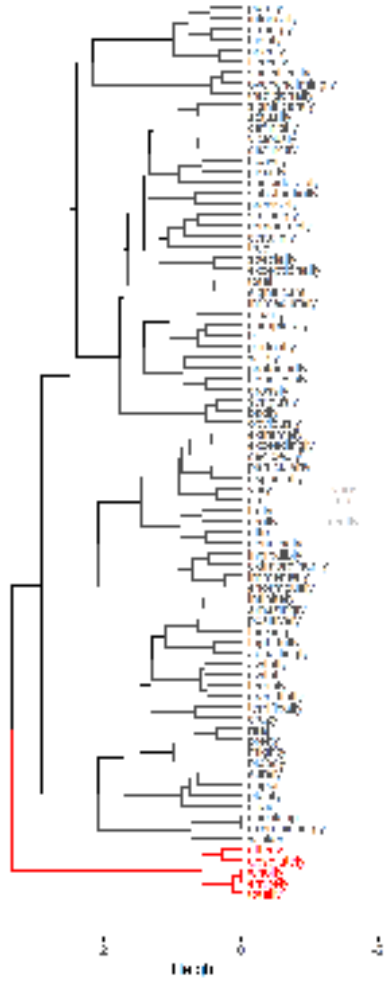


Figure 10: Semantic similarity of amplifier types in AmE (COHA) Schweinberger (fca).

Figure 11: Semantic similarity of amplifier types in AmE (COHA) Schweinberger (fca).

Discussion & Outlook

Summary

The analysis shows that . . .

- adjective amplification can show prolonged phases of stasis (no or little changes in relative frequency)
  → amplification is, at least not unconditionally, a domain of "[. . .] waxing and waning of forms" (D'Arcy 2015: 450)
- observed change is restricted to predicative contexts (thus making it unlikely that *so* replaces *very* long-term or across moralities (as is the case with *really* in informal spoken discourse)
- prolonged phases of stasis can be interrupted by sudden shifts which contradicts the notion that change is (always) gradual and step-wise
  → punctuated equilibrium!?!

Discussion

- *so* can successfully replace the dominant form (*very*) because both variants are semantically similar
  → replacement requires semantic similarity! (see also Schweinberger (fcb))
- No signs that *so* broadens before taking over the system.
- Moderate signs that *so* specialized during its increase (increase in collocation strength with *good* and *glad*)

Problems and Outlook

Problems

- Uneven distribution of letters across time periods (maybe use CORICOR for future research)
- Speaker-variability is not yet included into the models (run a mixed-effects model with speakers as random effect)
- No answer to the question as to what triggered the increase in *so* in both IrE and AmE...

Outlook

- Similar findings have been described for data from the fiction(!) section of the *Corpus of Historical American English* (COHA)
  $\rightarrow$ universal trend in "speech-like" writing or is it a contact phenomenon or something else?

Thank you so, really, very much!

Acknowledgements

I would like to thank...

Peter Siemund and his team (Lukas Pietsch, Kalynda Beal)
for providing me with a preliminary version of the HCIE
(without them the current study would not have been possible)

my colleagues at UQ
for comments and their feedback on earlier versions of this talk

This study will appear as:
Schweinberger, Martin. accepted. Analyzing Historical Changes in the Irish English Amplifier System. Anglistik – International Journal of English Studies (Special Issue: Focus on English Linguistics - Varieties meet Histories, Eds. Daniela Kolbe-Hannah & Ilse Wischer)

Aijmer, K. (2011). Are you totally spy? a new intensifier in present-day american english. In S. Hancil (Ed.), *Marqueurs discursifs et subjectivité*, pp. 155–172. Rouen: Universités de Rouen and Havre.

Aijmer, K. (2018). That's well bad. some new intensifiers in spoken in british english. In V. Brezina, R. Love, and K. Aijmer (Eds.), *Corpus Approaches to Contemporary British English*, pp. 60–95. New York and London: Routledge.

Barnfield, K. and I. Buchstaller (2010). Intensifiers on tyneside - longitudinal developments and new trends. *English World-Wide 31*(3), 252–287.

Bauer, L. and W. Bauer (2002). Adjective boosters in the english of young new zealanders. *Journal of English Linguistics 30*, 244–257.

Bolinger, D. (1972). *Degree Words*. The Hague: Mouton.

Borst, E. (1902). *Die Gradadverbien im Englischen*. Heidelberg: Winter.

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.

Fuchs, R. (2016). Register variation in intensifier usage across asian englishes. In H. Pichler (Ed.), *Discourse-Pragmatic Variation and Change: Insights from English*, pp. 185–213. Cambridge: Cambridge University Press.

Fuchs, R. (2017). Do women (still) use more intensifiers than men? *International Journal of Corpus Linguistics 22*(3), 345–374.

Hilpert, M. (2006). Keeping and eye on the data: Metonymies and their patterns. In A. Stefanowitsch and S. T. Gries (Eds.), *Corpus-Based Approaches to Metaphor and Metonymy*, pp. 123–151. Berlin & New York: Mouton de Gruyter.

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.

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 and G. Diewald (Eds.), *New Reflections on Grammaticalization*, pp. 143–161. Amsterdam: John Benjamins.

Macaulay, R. (2006). Pure grammaticalization: The development of a teenage intensifier. *Language Variation and Change 18*, 267–283.

Mair, C. (2004). Corpus linguistics and grammaticalisation theory. statistics, frequencies and beyond. In H. Lindquist and C. Mair (Eds.), *Corpus approaches to grammaticalization in English*, pp. 121–150. Amsterdam and Philadelphia: John Benjamins.

Méndez-Naya, B. (2003). On intensifiers and grammaticalization: The case of swithe. *English Studies 84*, 372–391.

Núñez Pertejo, P. and I. Palacios (2014). That's absolutely crap, totally rubbish. the use of intensifiers absolutely and totally in the spoken language of british adults and teenagers. *Functions of Language 21*(2), 210–237.

Palacios, I. and P. Núñez Pertejo (2012). He's absolutely massive. it's a super day. madonna, she is a wicked singer. youth language and intensification: A corpus-based study. *Text and Talk 32*(6), 773–796.

Partington, A. (1993). Corpus evidence of language change: The case of intensifiers. In M. Baker, F. Gill, and E. Tognini-Bonelli (Eds.), *Text and technology: In honour of John Sinclair*, pp. 177–192. Amsterdam & Philadelphia: John Benjamins.

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

Schweinberger, M. (fc.a). Analyzing change in the american english amplifier system in the fiction genre. In P. Rautionaho, A. Nurmi, and J. Klemola (Eds.), *Corpora and the Changing Society: Studies in the Evolution of English*, pp. —. John Benjamins.

Schweinberger, M. (fc.b). 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. —. Frankfurt, a.M.: Peter Lang.

Schweinberger, M. (subm.). Absolutely fantastic and really, really good – language variation and change in irish english. In C. Amador-Moreno and S. Lucek (Eds.), *Expanding the Landscapes of Irish English: Research in Honour of Jeffrey Kallen*, pp. —. John Benjamins.

Stefanowitsch, A. and S. T. Gries (2005). Covarying collexemes. *Corpus linguistics and linguistic theory 1*(1), 1–43.

Tagliamonte, S. (2006). "so cool, right?": Canadian english entering the 21st century. *The Canadian Journal of Linguistics/La revue canadienne de linguistique 51*(2), 309–331.

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

Tagliamonte, S. and C. Roberts (2005). So weird; so cool; so innovative: The use of intensifiers in the television series friends. *American Speech 80*(3), 280–300.

Tagliamonte, S. A. and D. Denis (2014). Expanding the transmission/diffusion dichotomy: Evidence from canada. *Language 90*(1), 90–136.

Analyzing historical changes in Irish English adjective amplification

slides available at

www.martinschweinberger.de

m.schweinberger@uq.edu.au

R code upon request

Dr. Martin Schweinberger

The University
Of Queensland
AUSTRALIA