Absolutely fantastic! Adjective amplification in English

Dr. Martin Schweinberger

slides available at

www.martinschweinberger.de

R code upon request

The University Of Queensland

AUSTRALIA

Introduction to Amplification

A Usage-Based Classification of Amplifiers

Amplification in L1-acquisition

Amplification in SLA

Amplification within and across varieties

Aims of this talk

- Tell you something about adjective amplification
- Introduce basic concepts of CxG/ Usage-Based approaches
- Exemplify how quantitative, statistical methods can be used fruitfully in Linguistics and SLA research

Introduction to Amplification

What is Amplification?

Phenomenon: Adjective Amplification

(1) And you just have to hint well then it's a very good hint (ICE-AUS:S1A-012$A)

(2) They're all really cheap <#> They're all really nice, the t-shirts in there (ICE-AUS:S1A-009$B)

(3) It was so bad (ICE-AUS:S1A-044$B)

Intensification and Amplification

Intensification is 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*

Forms of Amplification

(4) Lexical
very, *real(ly)*, *extremely*, *totally*, etc.

(5) Morphological
{uber#}, *{super#}*,*{hyper#}*,*{mega#}*,

Syntactic function of adjective

(6) Attributive
The very/so hungry caterpillar is nice.

(7) Predicative
The nice caterpillar is very/so hungry.

very vs. *really*: no meaning change → interchangeable
very vs. *hardly*: meaning change → not interchangeable

Motivation for Studying Amplification

Amplification

- Major area of gramm. change (cf. Brinton and Arnovick 2006: 441)
- Crucial for "social and emotional expression of speakers" (Ito and Tagliamonte 2003: 258)
- Linguistic subsystem which allows precise circumscription of a variable context (Labov 1972, 1966: 49)
- Ideal case for testing mechanisms underlying language change!

Previous research

- Extensive history of research on intensifiers (e.g. Borst 1902; Bolinger 1972)
- Intensification is considered a major area of grammatical change in English (cf. Brinton and Arnovick 2006: 441)
- Growing amount of variationist and historical research (e.g. Ito and Tagliamonte 2003; Tagliamonte and Roberts 2005; Macaulay 2006; Tagliamonte 2006)
- Very little research on the acquisition of intensification(!); an exception is Gülzow (2006)

(8) The queen herself welcomed the soldiers

A Usage-Based Classification of Amplifiers

Usage-Based/Data-Driven Classification

General idea

- Semantic Vector Space Modeling based on co-occurrence profiles of adjectives and amplifiers
- Amplifiers that co-occur with the same adjectives are semantically similar and thus interchangable

Advantages

- Relatively easy and can take variety specific differences and changes in use/meaning into account
- Does not rely on any theoretical framework
- Can provide a more fine-grained classification (different groups/clusters) that also informs about statistical significance

How data-driven classification works

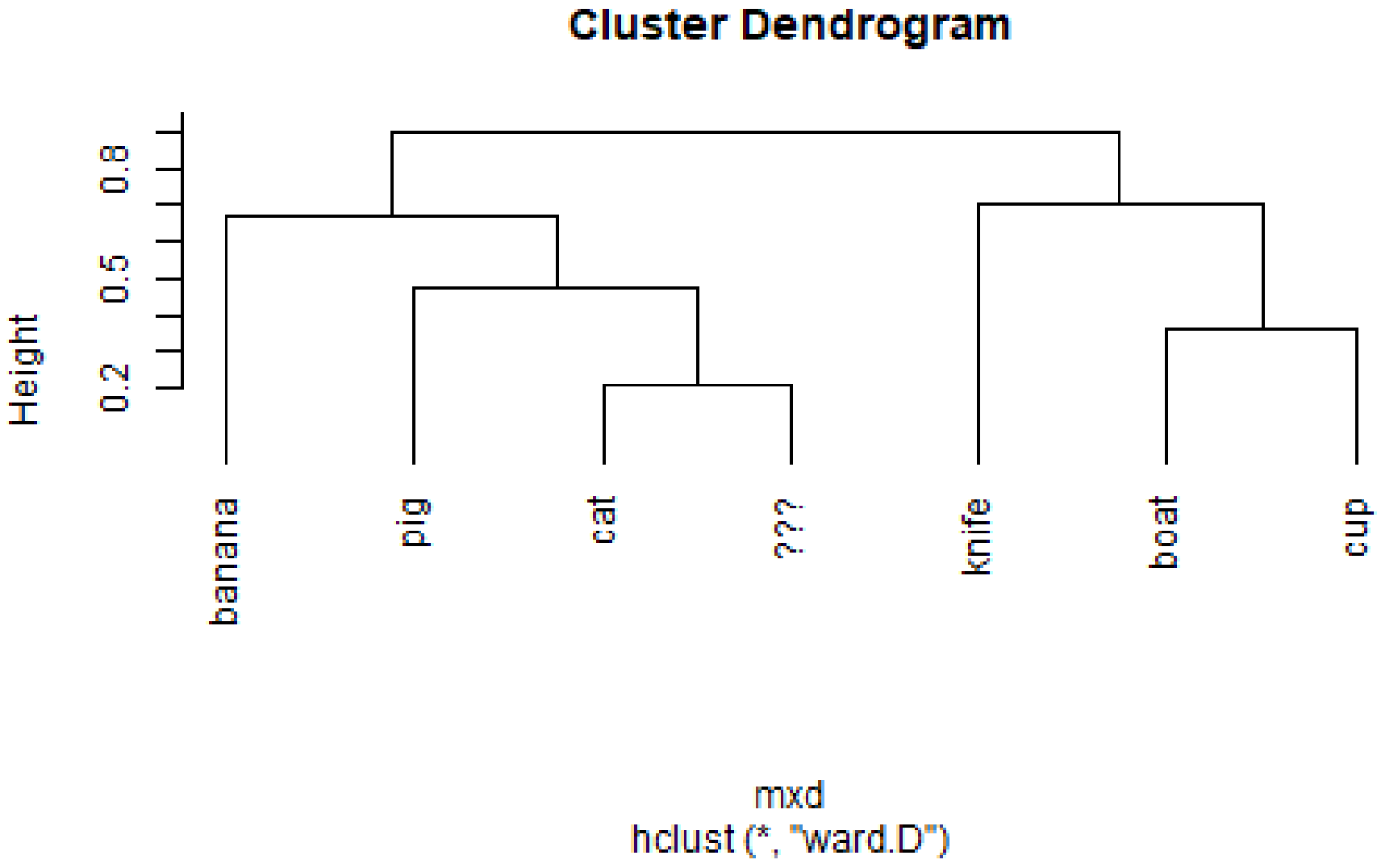
	get	**see**	**use**	**hear**	**eat**	**kill**
knife	31	16	69	0	2	0
cat	36	38	4	4	6	20
???	66	58	9	34	28	12
boat	46	21	17	4	0	0
cup	59	6	5	1	1	0
pig	4	15	3	1	7	21
banana	7	2	2	0	12	0

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

How data-driven classification works

	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")

Q_1

Does a data driven classification reflect the classification proposed in the literature?

$\rightarrow$ Are there meaningful clusters of amplifiers?

International Corpus of English (ICE)

- Australian ICE component
- One million words (600,000 spoken and 400,000 written) from diverse spoken and written text types (cf. next slide) with each file containing app. 2,000 words.
- Accompanied by metadata and biodata of speakers
- For the semantic vector space modeling → only spoken data!

Corpus data: International Corpus of English (ICE)

Mode	Conversation type	Register	Text type	Number of text files
SPOKEN (300)	Dialogues (180)	Private (100)	Face-to-face conversations	90
			Phonecalls	10
		Public (80)	Classroom Lessons	20
			Broadcast Discussions	20
			Broadcast Interviews	10
			Parliamentary Debates	10
			Legal cross-examinations	10
			Business Transactions	10
	Monologues (120)	Unscripted (70)	Spontaneous commentaries	20
			Unscripted Speeches	30
			Demonstrations	10
			Legal Presentations	10
		Scripted (50)	Broadcast News	20
			Broadcast Talks	20
			Non-broadcast Talks	10

Table 3: Schematic overview of the common design shared by all ICE components.

Data Processing

- POS-tagged all utterances
- Extraction of all adjectives and subsequently identifying adj. preceded by an amplifier
- Removed
  - Adj. that did not occur before a full stop or before a noun (to determine syn. function; predicative vs attributive)
  - Adj. that were not intensified by at least two different amplifier types (e.g. *right honorable*)
  - Adj. that were preceded by downtoners
  - Adj. preceded by strange forms (e.g. *much*)
  - Adj. that were negated (e.g. *not good* or *not very nice*)
  - Comparative and superlative forms (e.g. *better* or *strongest*)

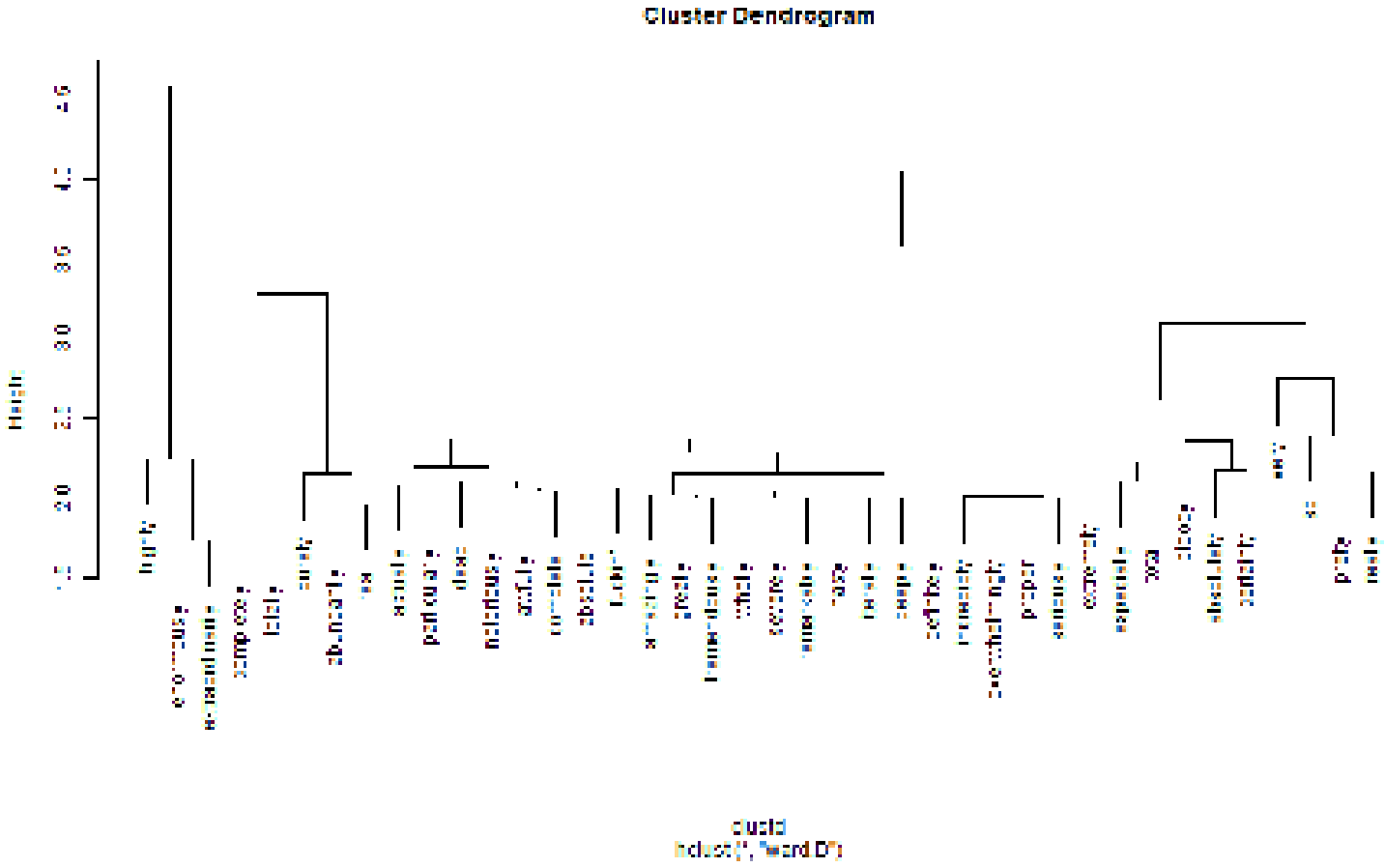


Figure 1: Rooted dendrogram showing the clustering of amplifiers in Australian English based on the semantic vector space model.

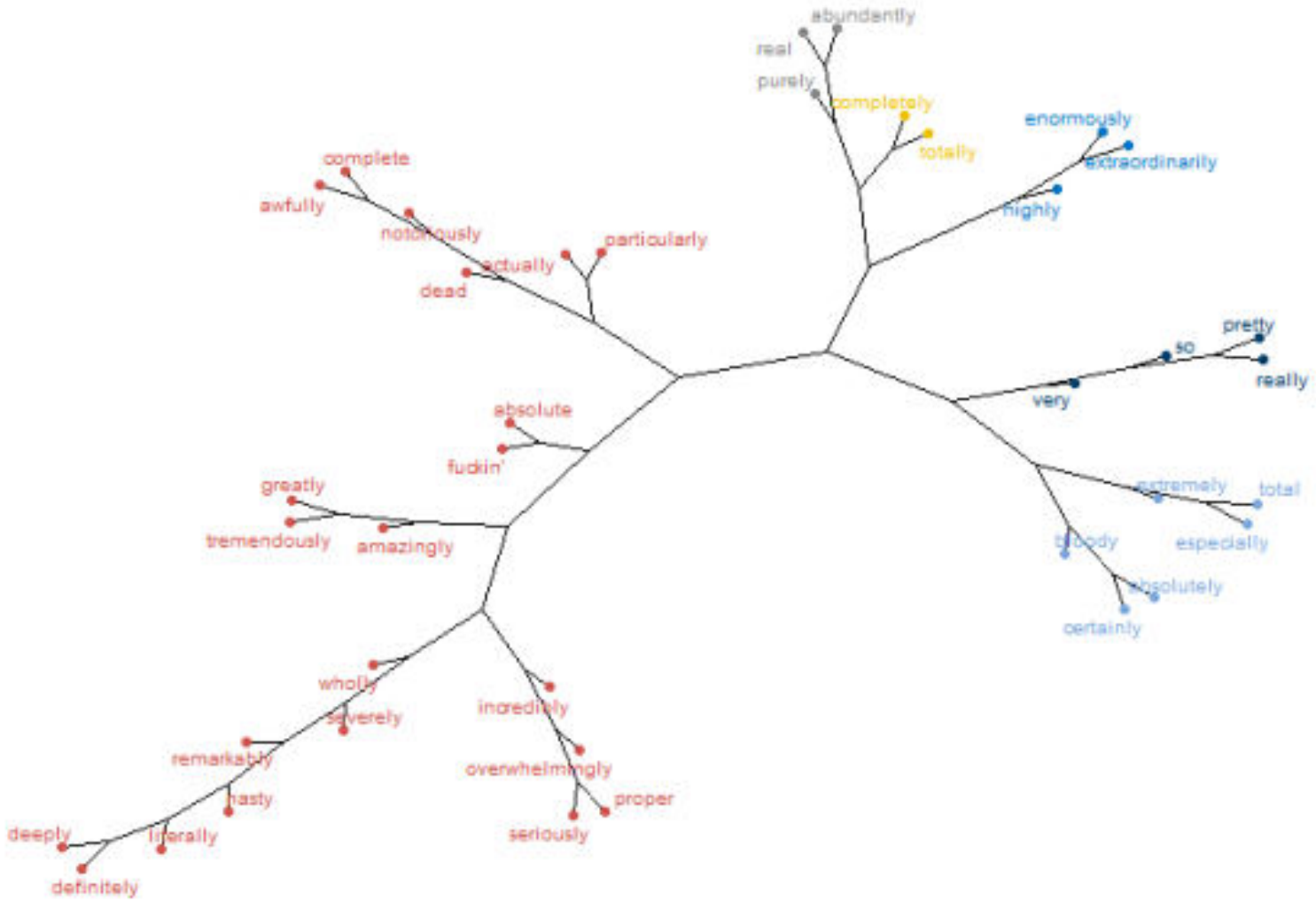


Figure 2: Unrooted dendrogram showing the clustering of amplifiers in Australian English based on the semantic vector space model.

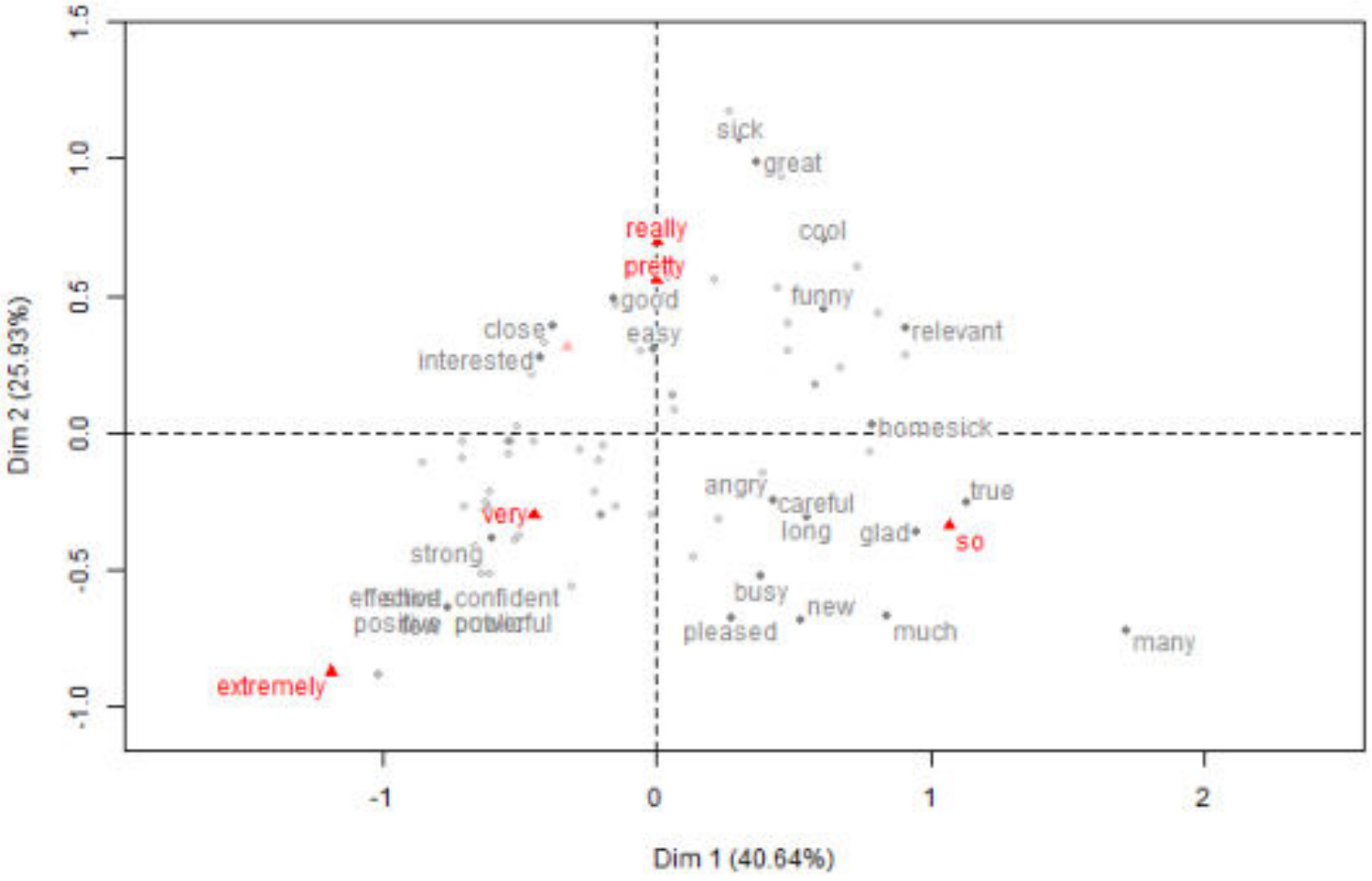


Figure 3: Results of a correspondence analysis based on amplifiers and their co-occurrences with adjectives in Australian English.

Amplification in L1-acquisition

Q_2

How do children acquire amplification?

$\rightarrow$ How does the use of amplifiers pattern among children?

Data

Home–School Study of Language and Literacy Development (part of CHILDES: *Child Language Data Exchange System*)

- Longitudinal data of English-speaking children from low-income families growing up in the Boston area.
- Transcripts collected during 5 home visits
- Visits took place at ages 3 (hv1), 4 (hv2), 5 (hv3), 2nd grade (hv4) and 4th grade (hv5)
- During visits children performed different tasks: book reading, toy play, child narratives, elicited report, and experimental tasks.

Data processing as described before.

Data Summary: HSLLD (Children only)

Amplifier	**N**	**%**	**Amp.** (%)
∅	4,776	95.16	
so	77	1.53	31.69
very	63	1.26	25.93
real	34	0.68	13.99
pretty	27	0.54	11.11
really	24	0.48	9.88
wicked	9	0.18	3.70
totally	5	0.10	2.06
completely	2	0.04	0.82
extra	1	0.02	0.41
fucking	1	0.02	0.41
Total	5,019	4.84	100

Table 4: Frequencies and percentages of amplifiers in variable contexts in the HSDDL.

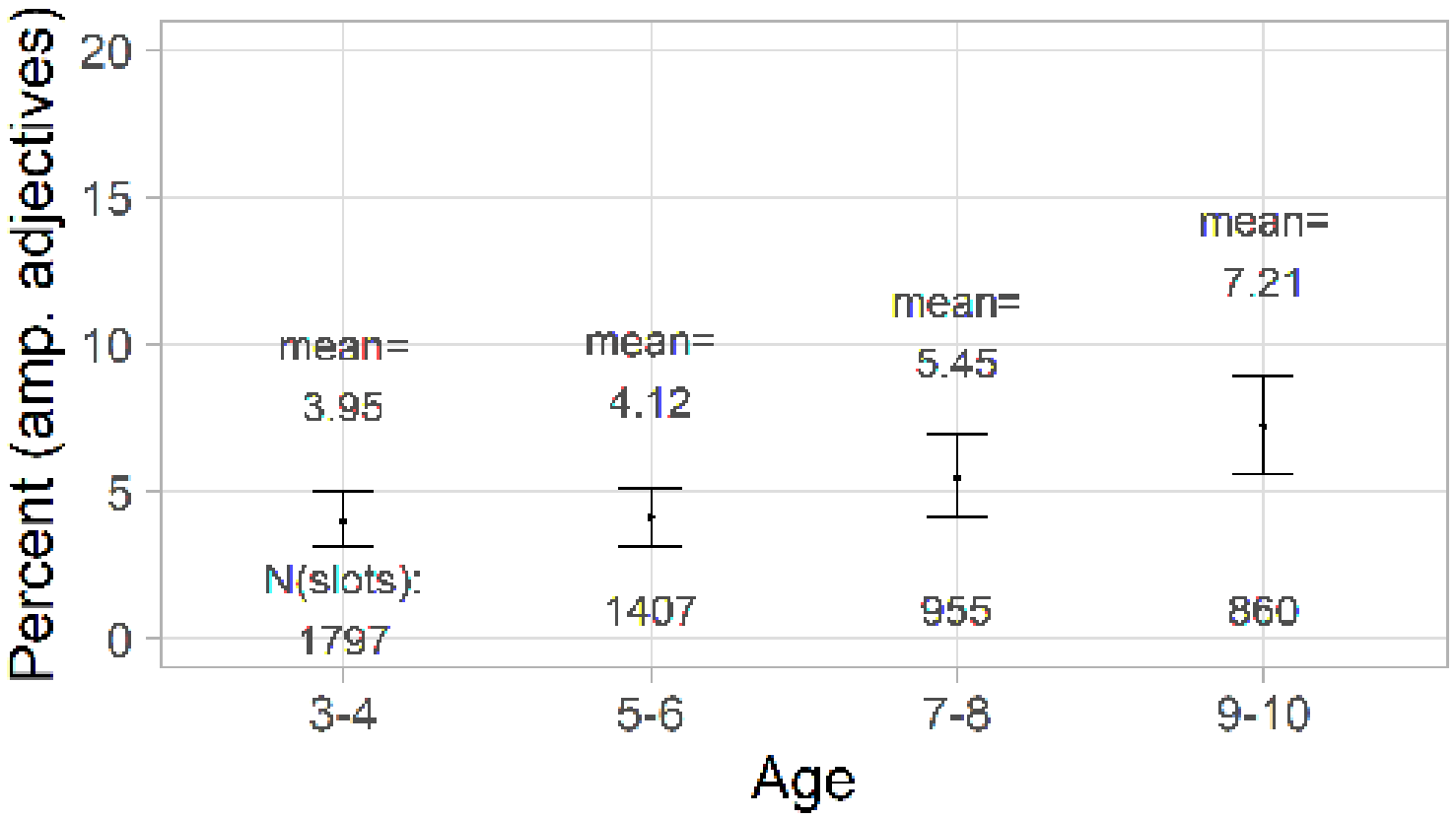


Figure 4: Percent of amplified adjectives by age (HSLLD)

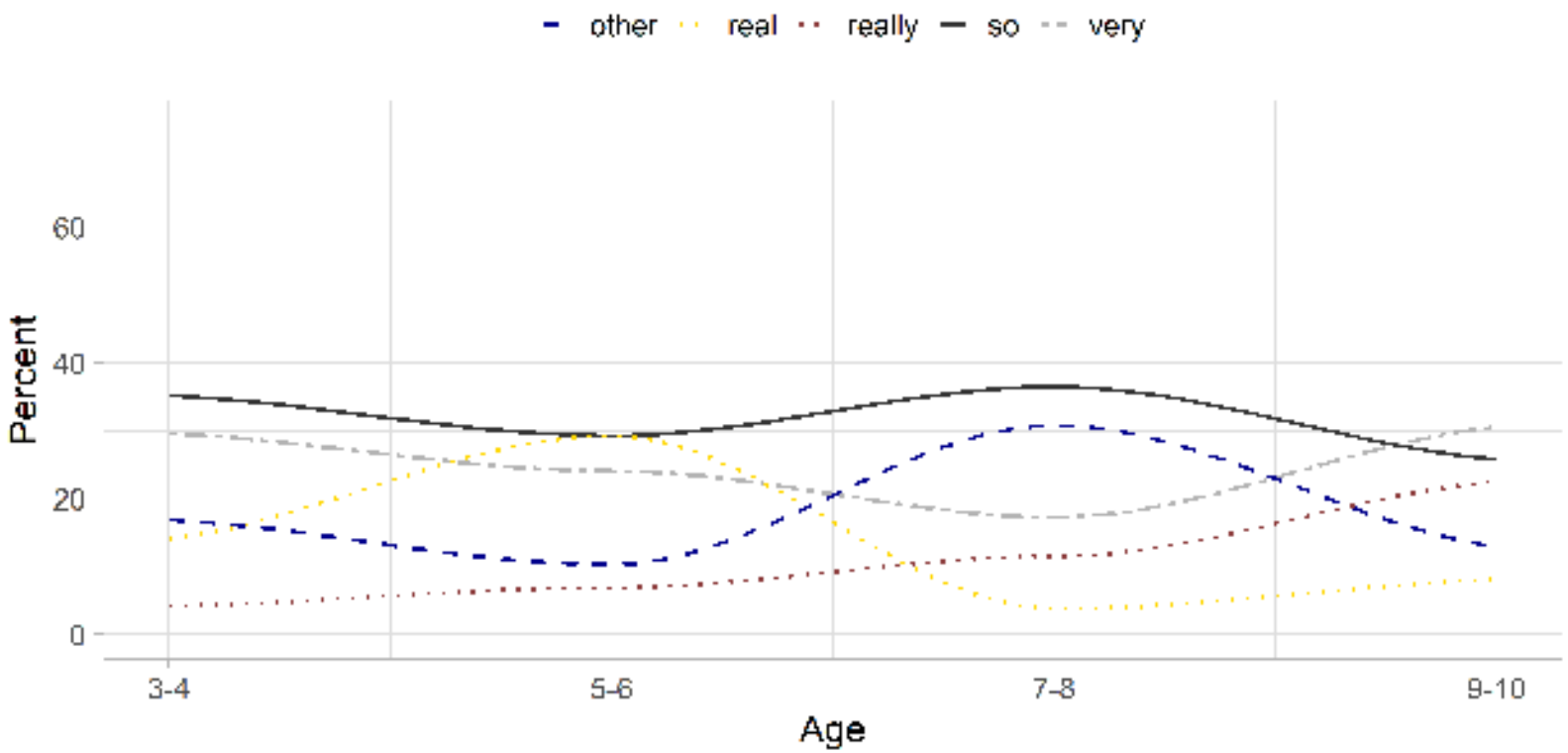


Figure 5: Amplifier types by age (HSLLD)

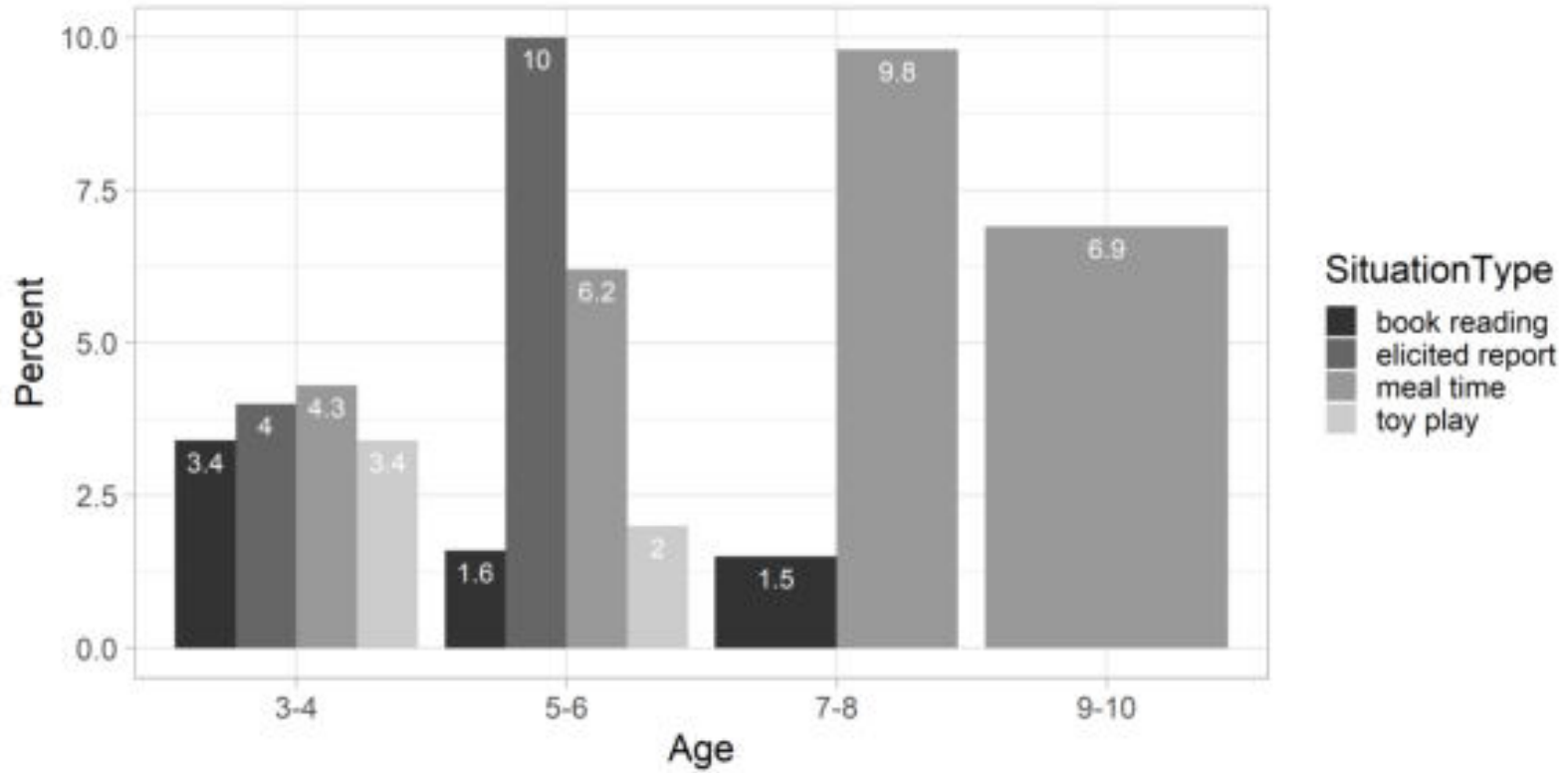


Figure 6: Percentages of amplification by age and situation type (HSLLD)

Interim synopsis

- Overall frequency
  Stability among younger speakers, increase in amplification among children aged 9 and 10
- Type frequency
  Chaotic use among younger children, patterning emerges after an age of 5 (*so* outperforms rival variants)
- Extra-ling. constraints
  - Similarity among children aged 3 and 4, situational differentiation emerges at age 5
  - Differentiation in use: freq. increase in meal time conversations but substantial decrease in book reading situations.

Q_3

How does the input of the mother pattern?

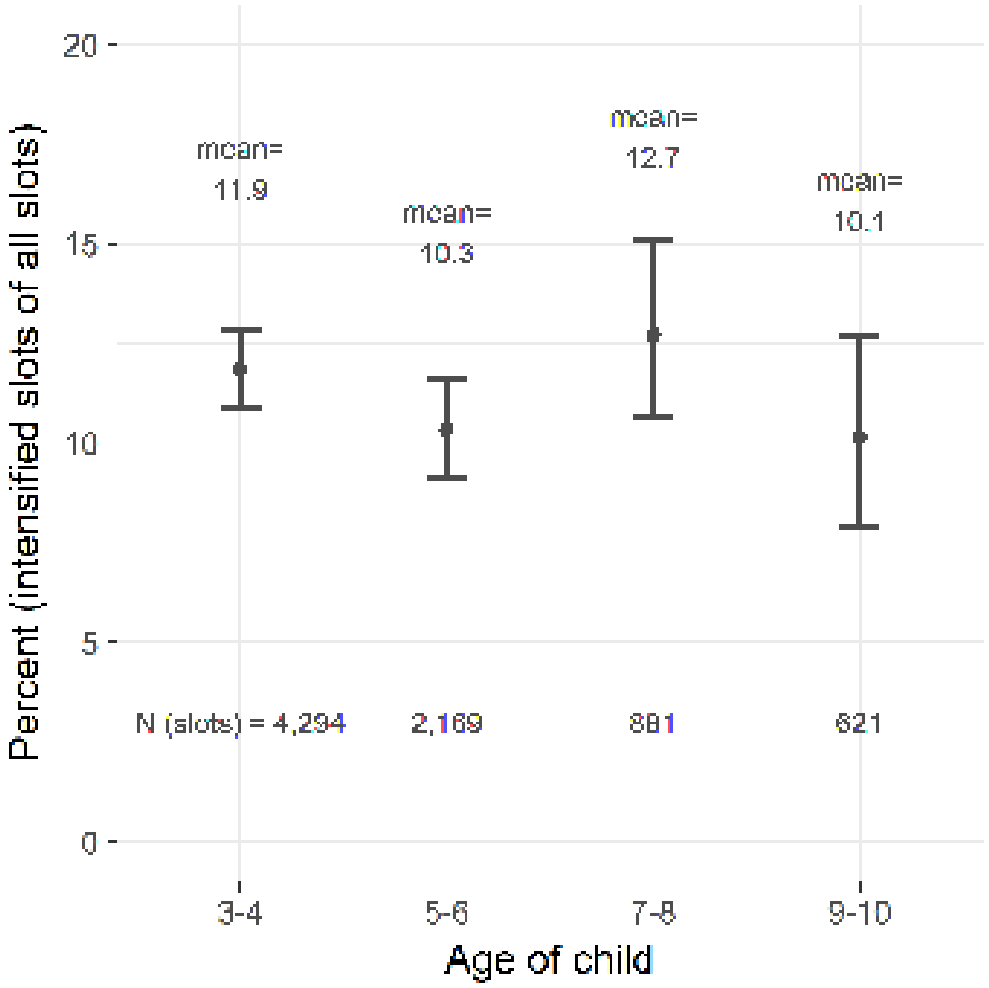


Figure 7: Percent of amplified slots in mother's CDS by age of child (HSLLD)

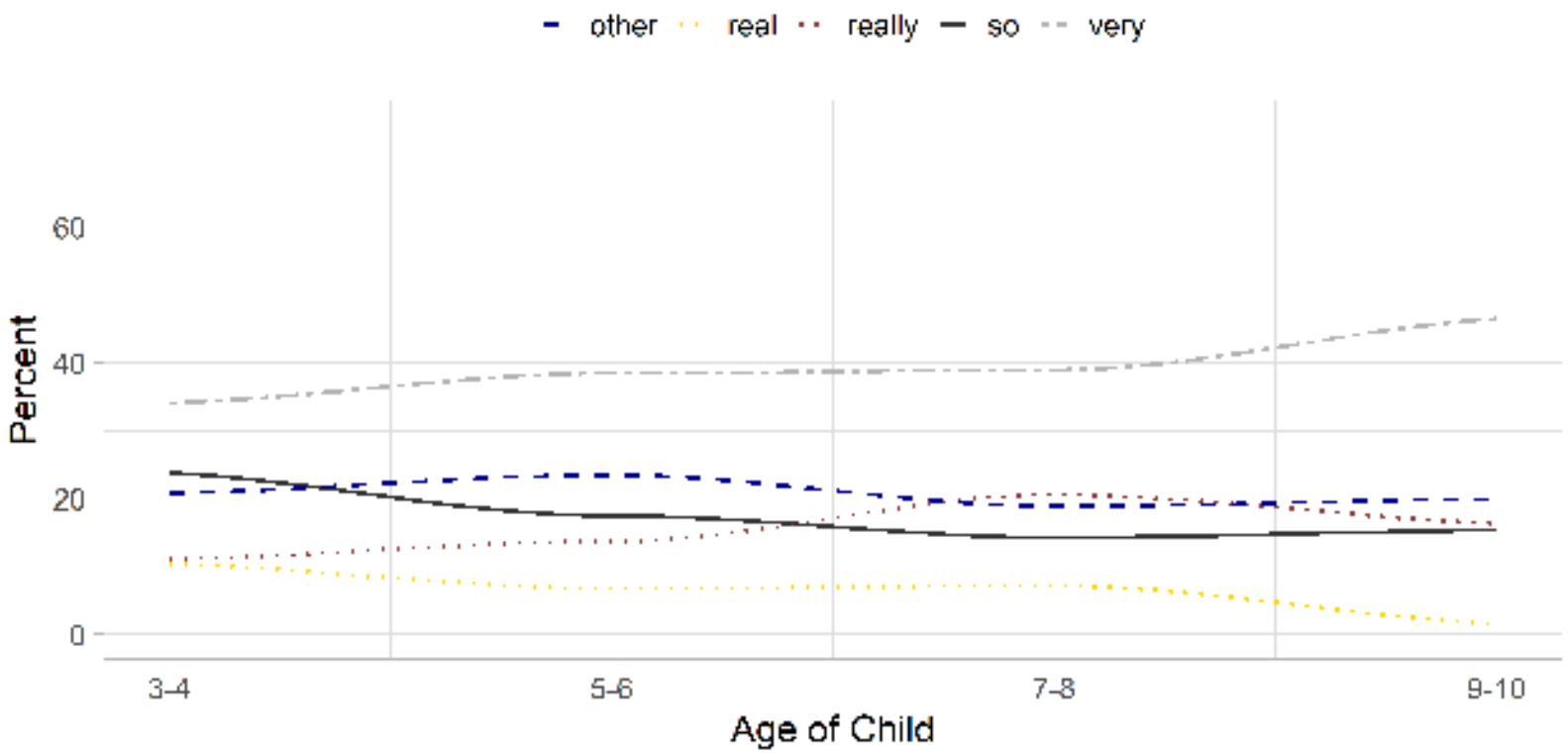


Figure 8: Mothers' amplifier types by child's age (HSLLD)

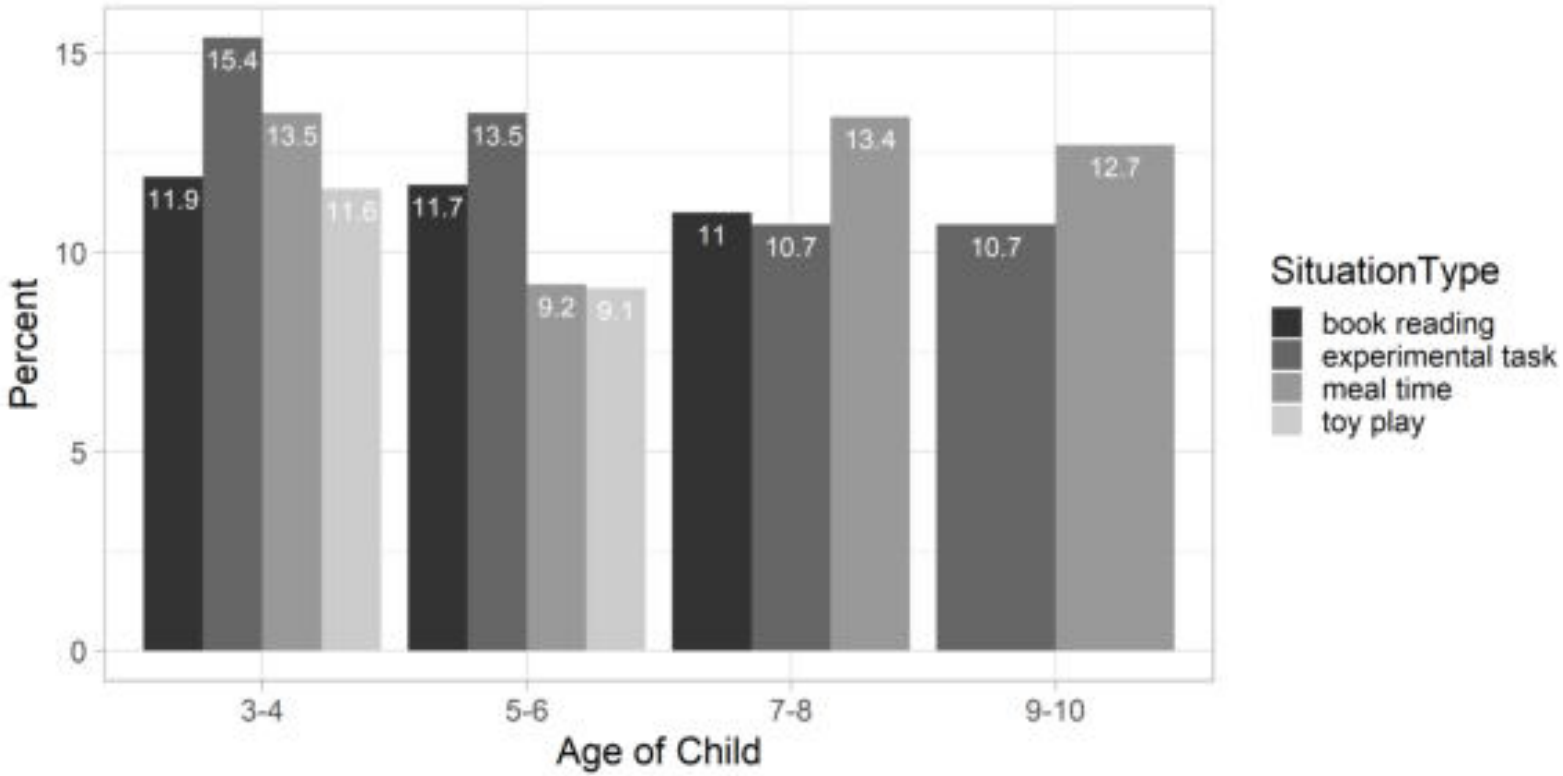


Figure 9: Percentages of amplification in mother's CDS by child's age and situation type (HSLLD)

Q_4

What causes the obs. stratification if not the CDS input?

$\rightarrow$ How does the use of women in non-CDS pattern?

Data and Data Processing

Santa Barbara Corpus of Spoken American English (SBC) (part of the American component of the *International Corpus of English* compiled between 2000 and 2005)

- Same processing as for HSLLD data
- Only speech of women between 19 and 50 years of age

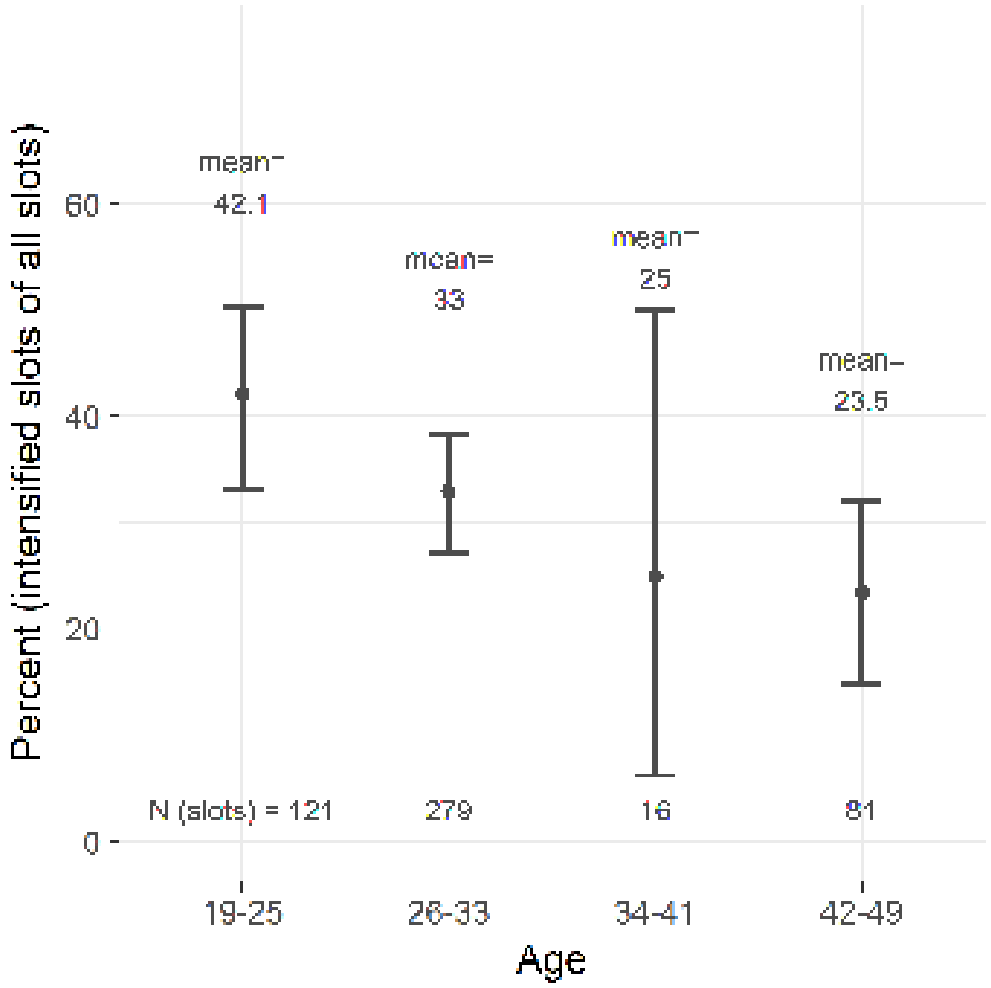


Figure 10: Percent of amplified adjectives by women against age (SBC)

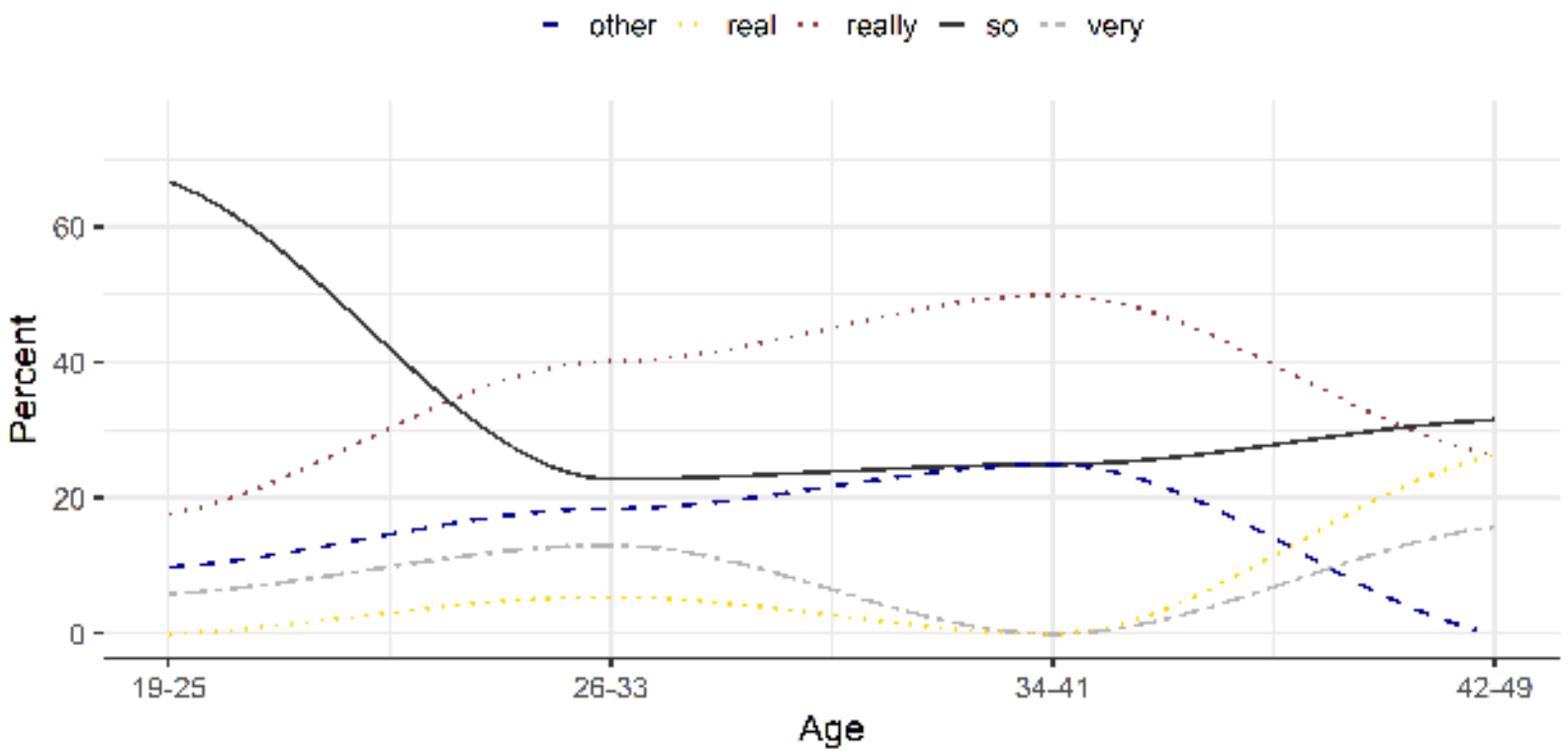
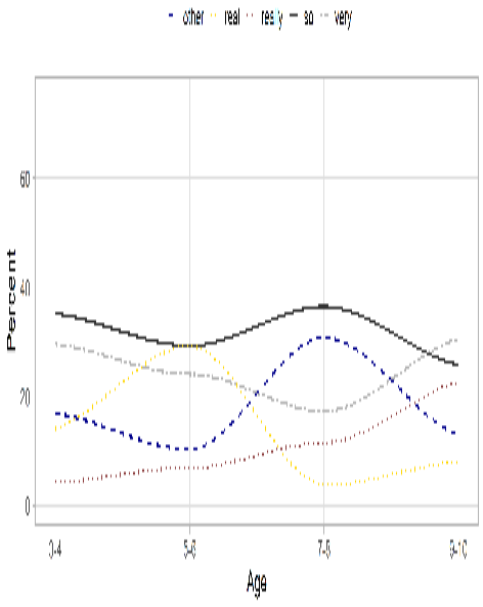
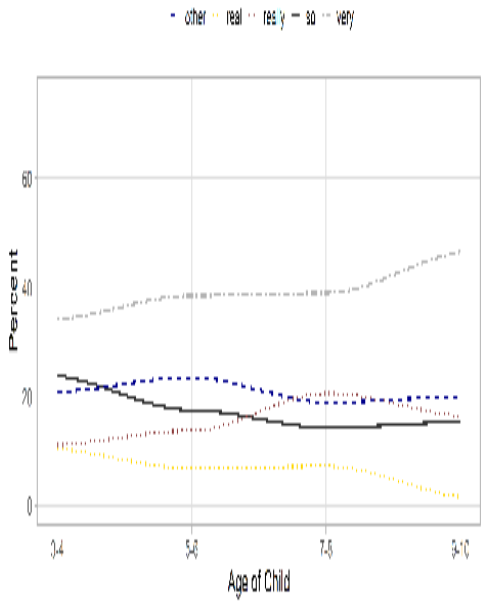


Figure 11: Women's amplifier types by age (SBC)

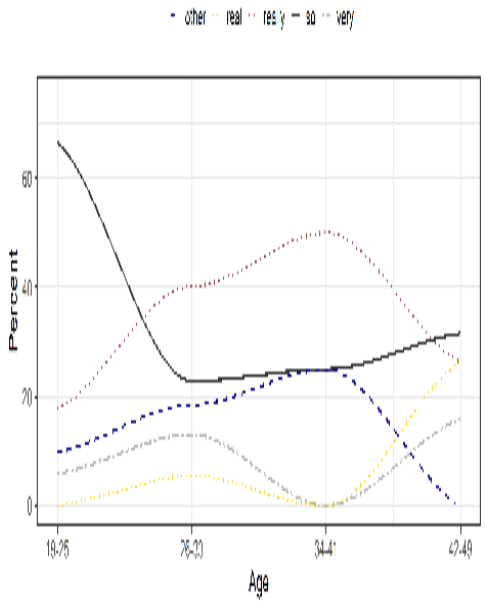
Interim recapitulation



Children



Mothers



Women

Interim synopsis

- Overall frequency
  Substantially higher compared to both children and CDS
- Type frequency
  Youngest group similar patterning to children
  (*so* > *really* > *very*)
- Extra-ling. constraints
  SBC represents private dialogue data: cannot test for effects of situation type (register)

Summary, Discussion & Problems

Summary

Main points (with pinches of salt)

- Frequency of children's use of amplifiers mirrors the frequency of CDS but not that of women of child bearing age (frequencies of youngest cohort in SBC much higher)
- Patterning of children's use of amplifier types mirrors the use of women of child bearing age (youngest cohort in SBC) but not CDS
- CDS very conservative(!) in terms of amplifier use
- Little register stratification in the speech of children aged 3-4 (similar to CDS)
- Drastic register differences are observable from age 5 onward (extra-ling. constraints!)

Discussion

- Results indicate that the children model their use based on non-CDS input rather than CDS (true for type patterning but not with respect to frequency).
- Frequency of amplification differs across situation types
  - extra-ling. constraints are acquired very early on.
  - extra-linguistic constraints seem to be acquired alongside linguistic forms rather than separate from linguistic forms as previously suggested (Labov 1964; Nardy et al. 2013: 258-260).

Discussion

L1-Acquisition from a usage-based CxG perspective

- Children are thought to start out with concrete pieces of language and gradually develop more schematic constructions. . . . Constructivists see these early constructions as the building blocks for later development. . . (Lieven 2006: 84–85)

Stage 1	**Stage 2**	**Stage 3**
Holophrases	Pivot schemas	Schematic constructions
↓	↓	↓
It's daddy! It's mommy! It's Elmo!	It's X_{NN}	$[Y_{Dem./Existential/Dummy} + X_{NN}]$
very busy very wet very hungry	very $X_{Adjective}$	$[Y_{Intensifier} + X_{Adjective}]_{Int.Construction}$

Discussion

Social Grounding of L1-Acquisition

- Children "acquire language in a socially grounded fashion. On the constructional view, the item-based schemas that children acquire are . . . tied to specific situations and situation types" (Hilpert 2014: 159)
- The patterning of amplifiers produced by children aged 5 and older supports predictions of usage-based approaches of language acquisition (cf. Tomasello 2003)

Remaining Issues

- Data set too small to warrant reliable/conclusive conclusions
- Situation types not across all age groups (data compilation not optimal; D'Arcy is compiling better data)
- SBC and HSLLD not fully comparable (SBC compiled later (2000-2005) than HSLLD (1987-1991) and regional difference: California vs Boston)

Amplification in SLA

Q_5

How and where do learners of English (NNS) differ from native speakers (NS) with respect to adjective amplification?

Data

- *International Corpus of Learners of English* (ICLE)
  - 2.5 mil. words representing argumentative writing by intermediate to advanced Bulgarian, Czech, Dutch, Finnish, French, German, Italian, Japanese, Norwegian, Polish, Russian, Spanish, and Swedish learners of English
- *Louvain Corpus of Native English Essays* (LOCNESS)
  - 290.000 words of argumentative essays by American and British university students and British A-level students
  - LOCNESS was specifically designed to allow meaningful comparisons between the learner data represented in the ICLE.

- Processing as described above.

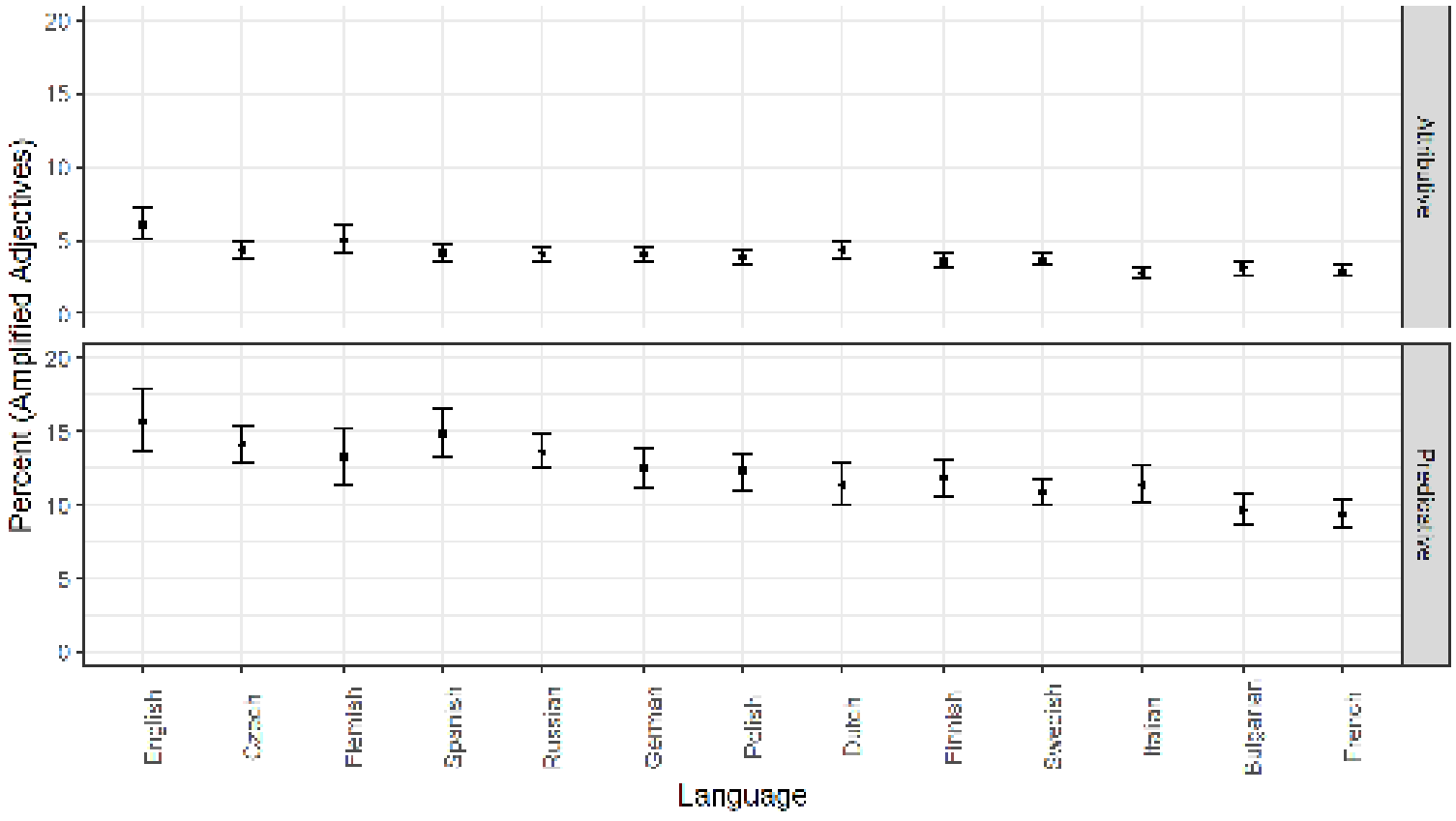


Figure 15: Percent of amplified slots across L1-backgrounds by syn. function in descending order.

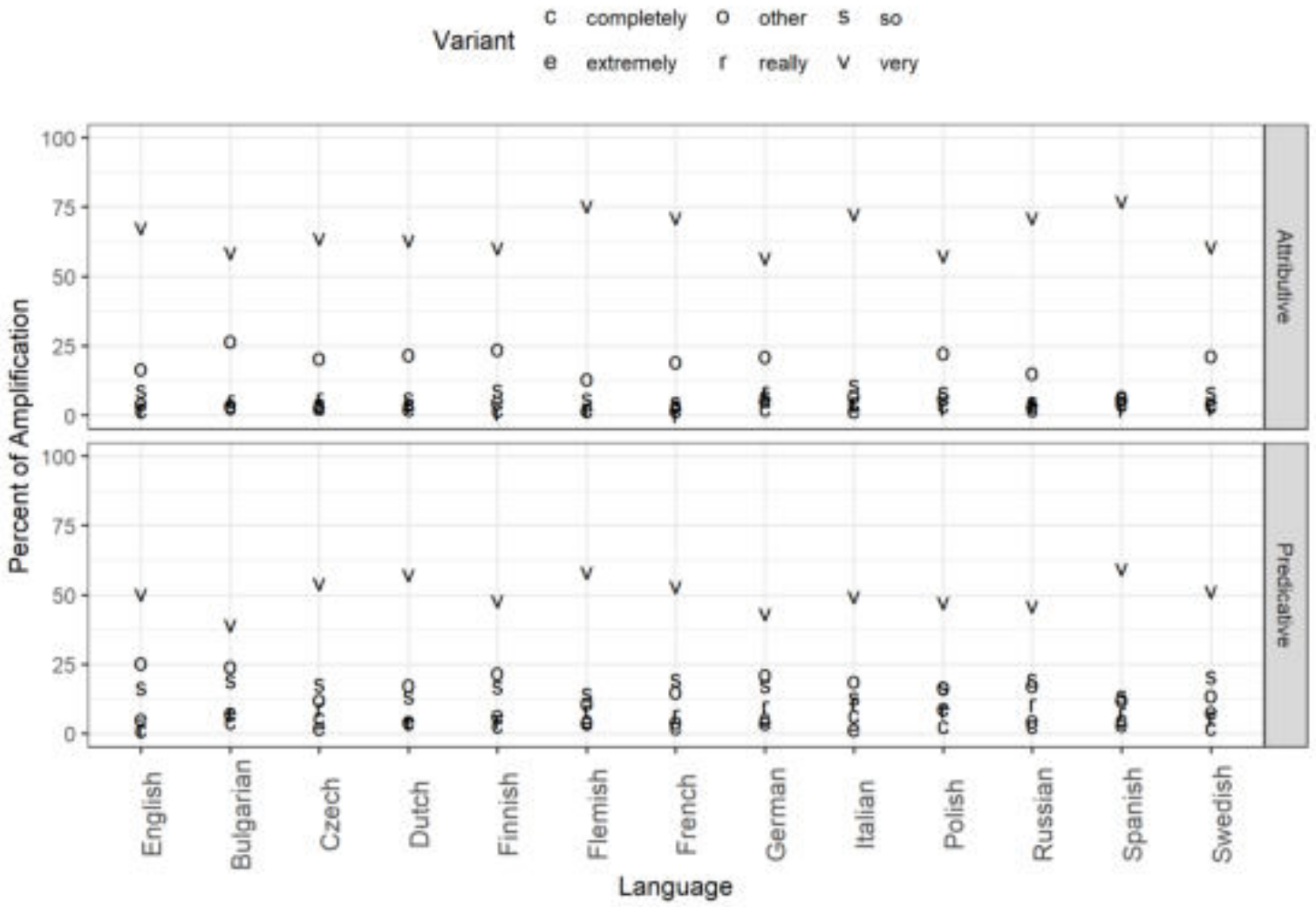


Figure 16: Percentages of amplifiers by L1-background.

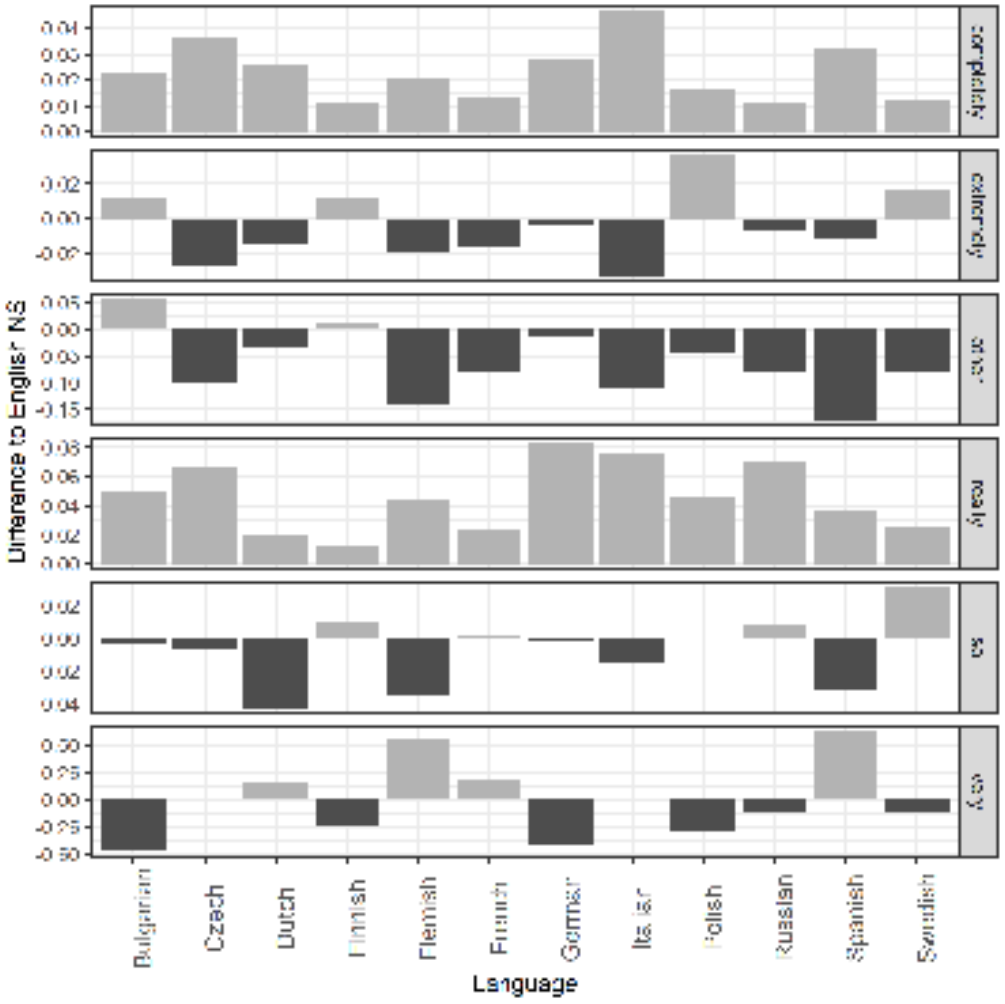


Figure 17: Bar graphs showing the difference to expected frequencies of amplifier types based on NS use.

Statistical Analysis

Covarying Collexeme Analysis (CCA) (Stefanowitsch and Gries 2005)

- Extension of Fisher's Exact test
- Evaluate attraction between elements that occur in two distinct slots within a specified construction
- How does a variant in a first slot affect the likelihood of another variant from another set in a second slot?
- Values below 0 indicate rejection while values above 0 indicate attraction
- Advantage of CCA
  - very robust(!)
  - does not rely on distributional assumptions (unlike the χ^2-family of tests)

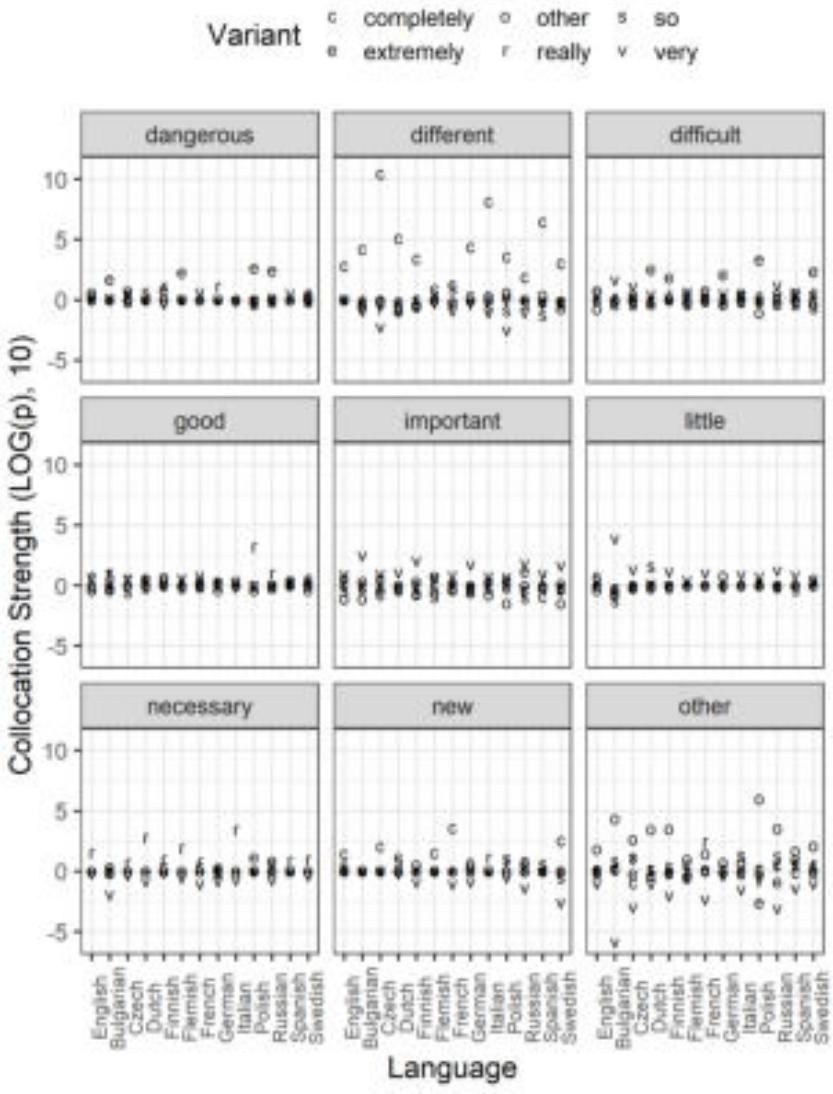


Figure 18: Results of the covarying collexeme analysis by L1-background and adjective.

Summary

- NS amplify more than NNS
- Patterning of amplification varies substantially across L1-backgrounds
- Common trends among NNS
  - Overuse of *completely* and *really*
  - Overuse of *completely* with *different*
  - Overuse of *really* with *necessary*
  - Overuse of *extremely* with *difficult*

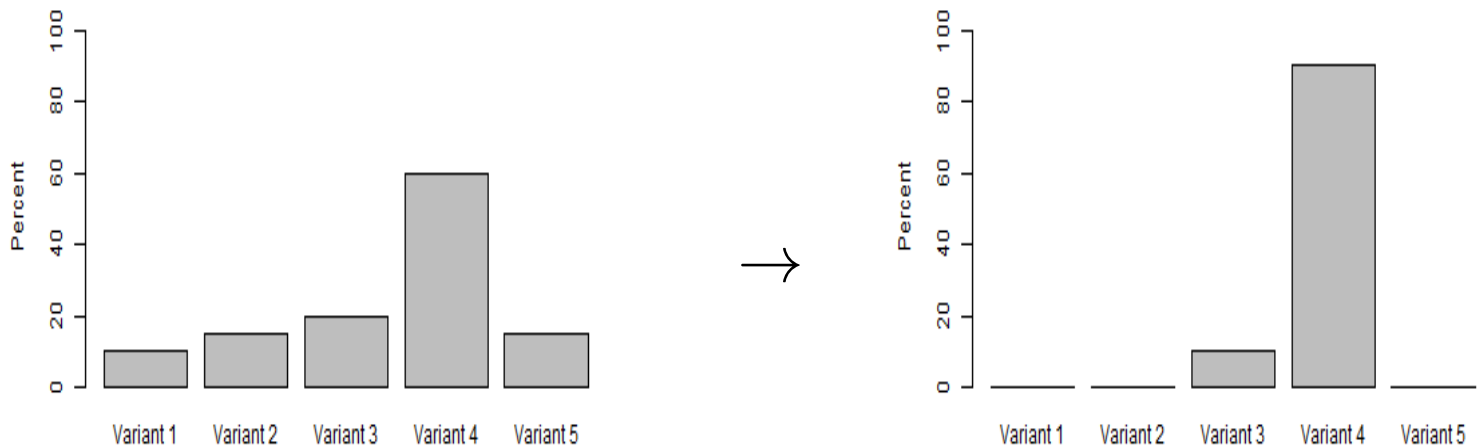
Discussion

- Results indicate that the NNS use collocational patterns that are typical of informal speech
- Lack of pragmatic/stylistic awareness (poverty of input)
- Language teachers can profit from learning about such systematic divergences among NNS
- NNS can profit from learning about stylistic constraints (pragmatic competence)

Discussion

Statistical Learning and SLA from a usage-based CxG perspective

- Children overgeneralize and form default Cxs while ignoring (systematic) variability/variation.



Input (by mother/later: peers

Output (by child)

Discussion

Statistical Learning and SLA from a usage-based CxG perspective

- Children overgeneralize and form default Cxs while ignoring (systematic) variability/variation.

Stage 1 Beginner (Holophrases)	**Stage 2** Advanced learner (Overgeneralization)	**Stage 3** Near native-like (Stylistic variability)
very good	really good	really good
very difficult	extremely difficult	very difficult
very different	completely different	completely different
very $X_{Adjective}$	fixed $[Y_{Intensifier}+X_{Adjective}]$	$[Y_{Intensifier}] + [X_{Adjective}]$

Amplification within and across varieties

Previous Research

Amplification

- substantial amount of corpus-based research on intensification (e.g Aijmer 2011, 2018; Fuchs 2016, 2017; Núñez Pertejo and Palacios 2014; Palacios and Núñez Pertejo 2012)
  → but mostly either focused on individual intensifiers or without regard to the intensified adjectives
- recently amplifier-adjective bigrams have come more into focus (e.g. Schweinberger 2017; Wagner 2017a,b)
- associated with teenage talk and young(ish) (female) speakers
  (Bauer and Bauer 2002; D'Arcy 2015; Macaulay 2006; Tagliamonte 2006, 2008)

Focus

- 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)

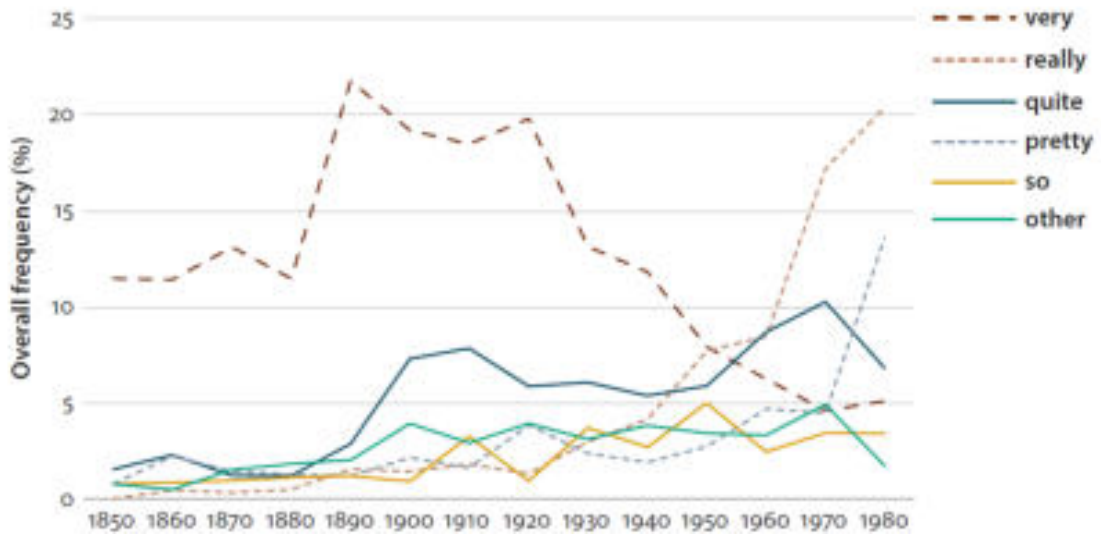


Figure 21: Adapted from D'Arcy (2015: 468)

Variationist Sociolinguistics

- Language is not homogeneous: variation is ubiquitous
  - Social factors : language use
  - Linguistic variation not random
  - Systematic correlation between certain social factors (age, gender, class, ethnicity, etc.) and language use
- Linguistic differentiation $\leftrightarrow$ social stratification

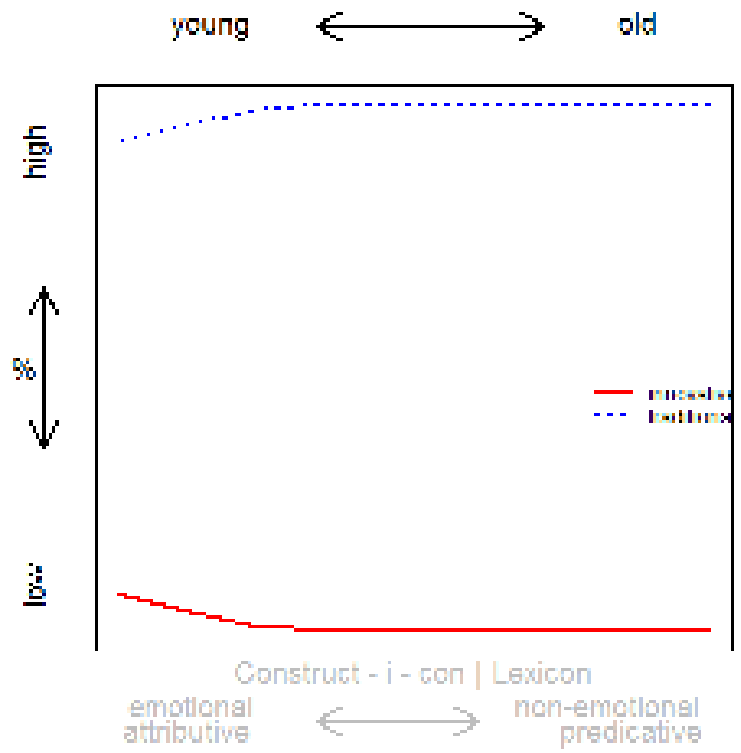
Diffusion of Innovations

young ⟷ old

upper

⟷

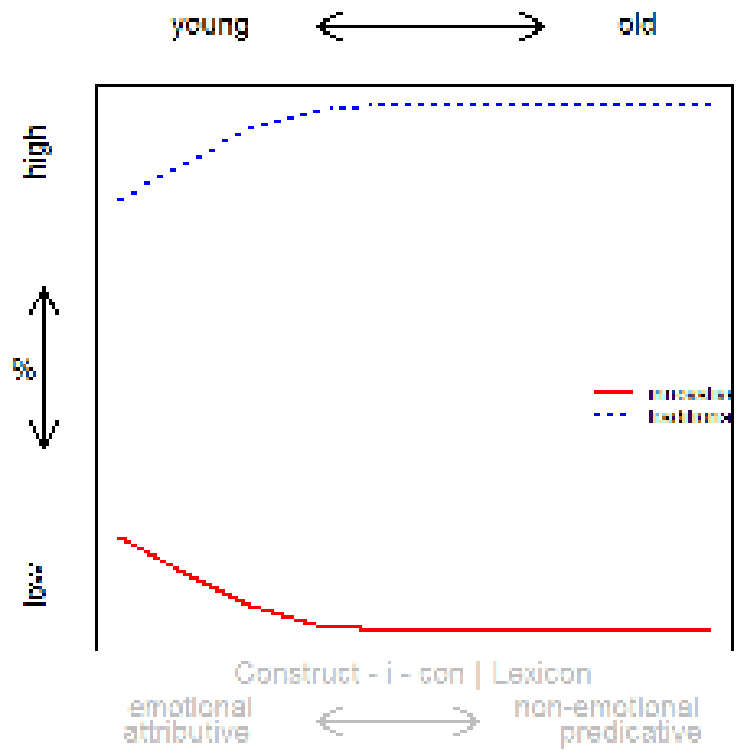
working



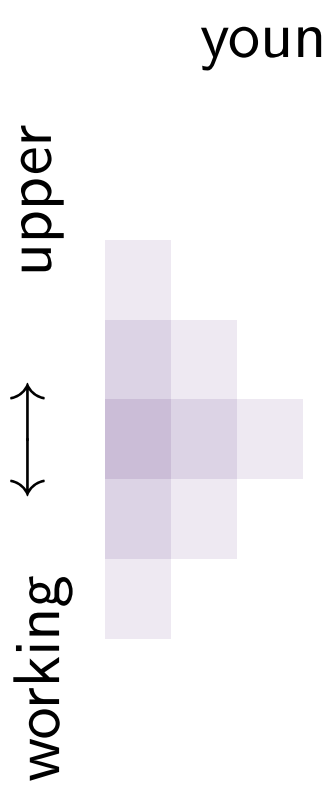
Diffusion of Innovations

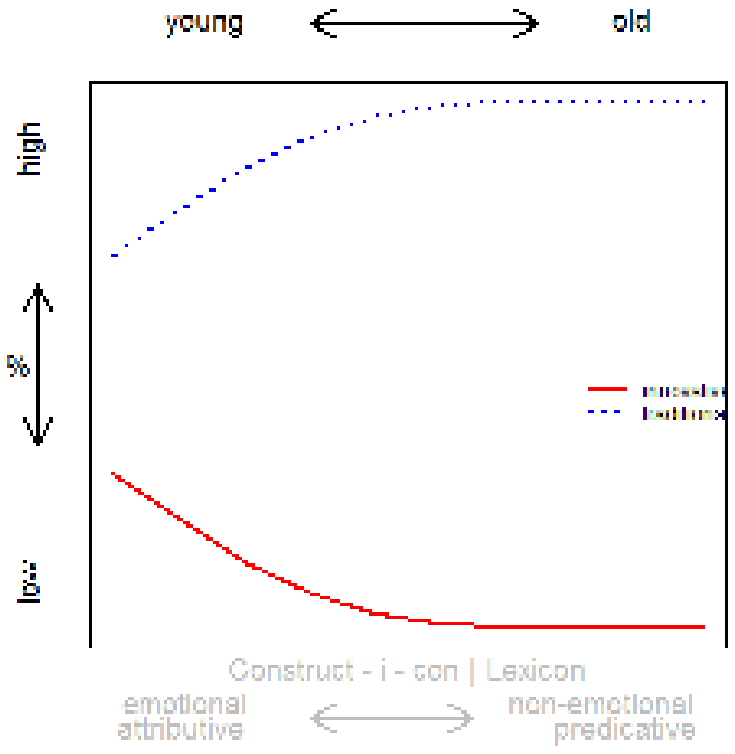
young ⟷ old

upper ⟷ working

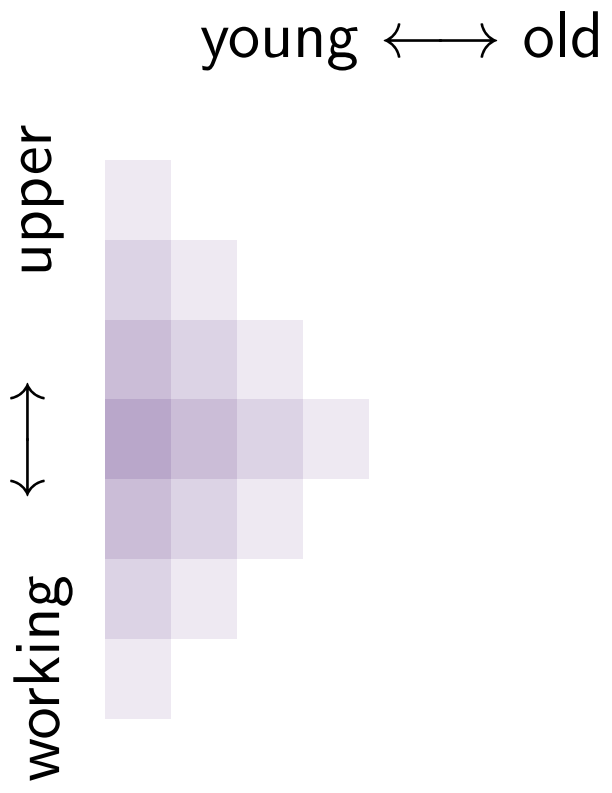


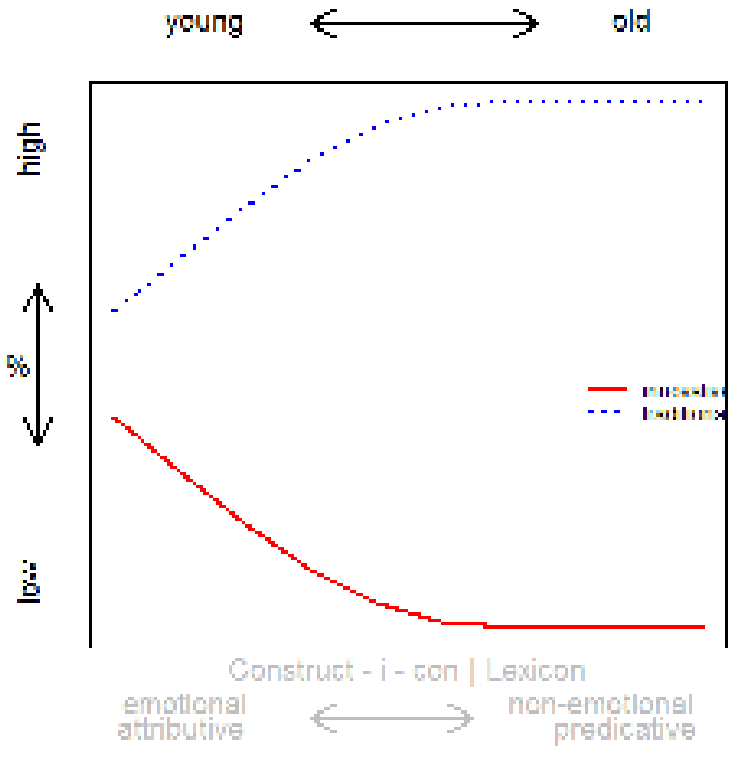
Diffusion of Innovations



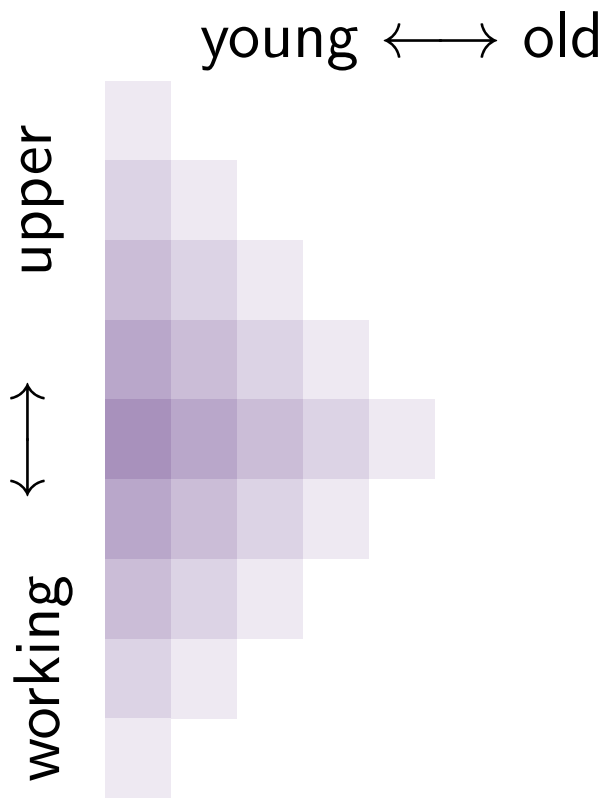


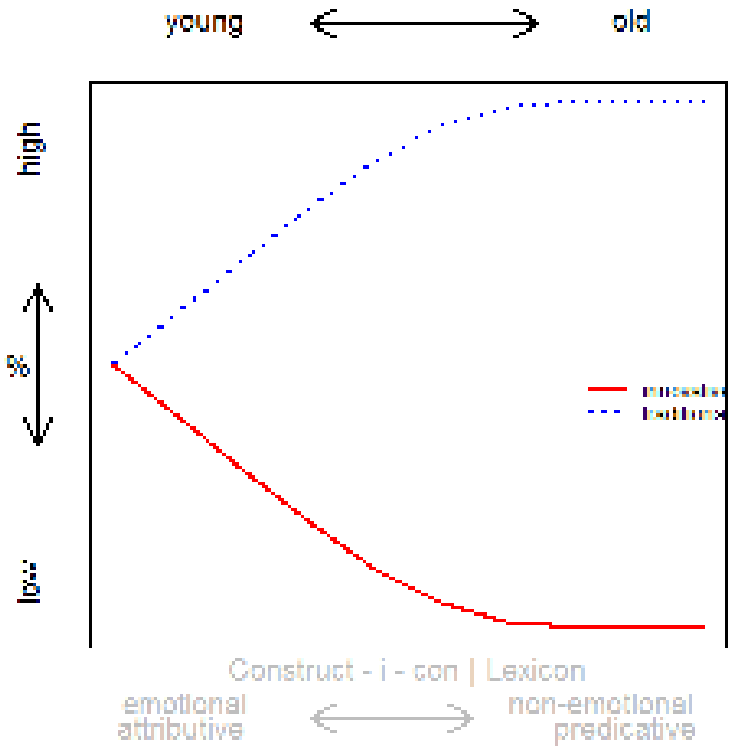
Diffusion of Innovations



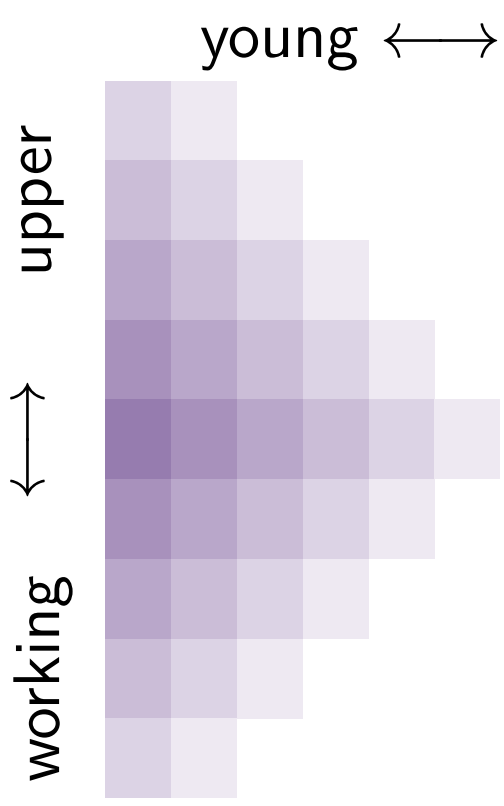


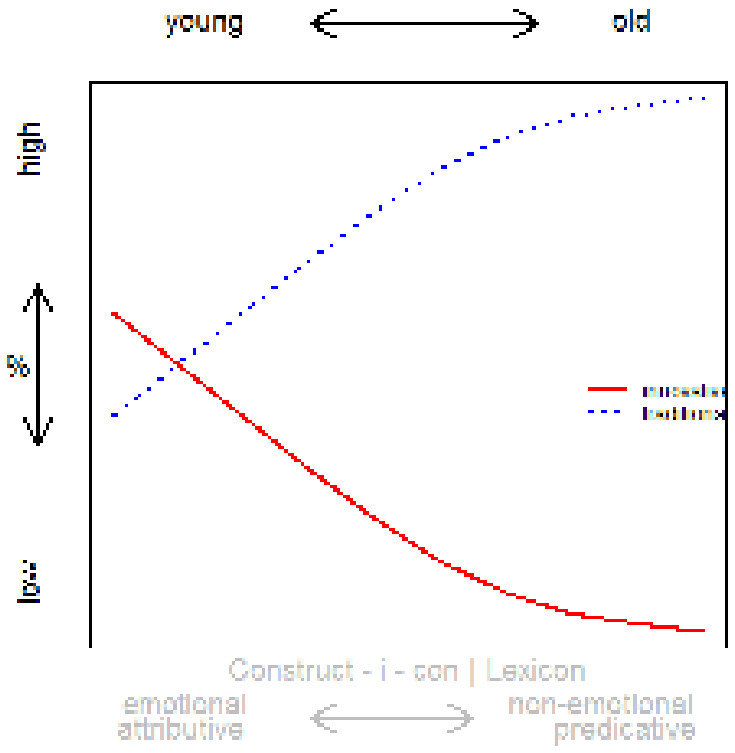
Diffusion of Innovations



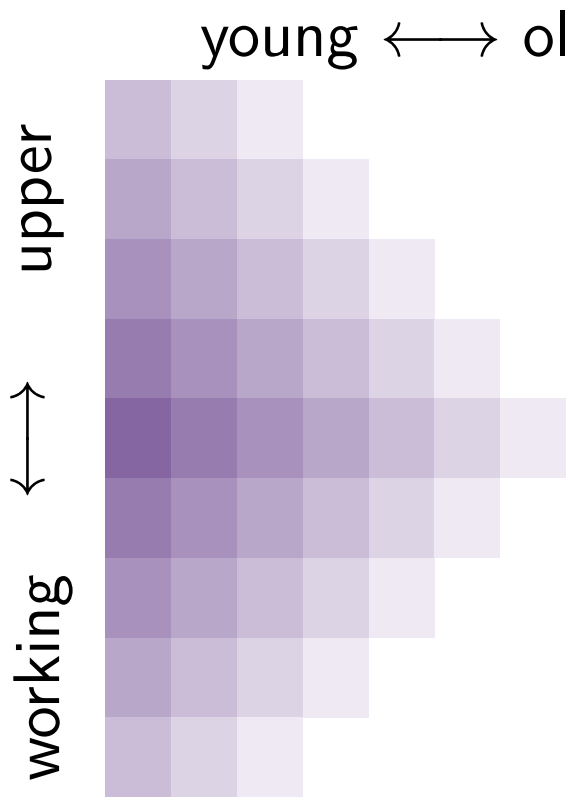


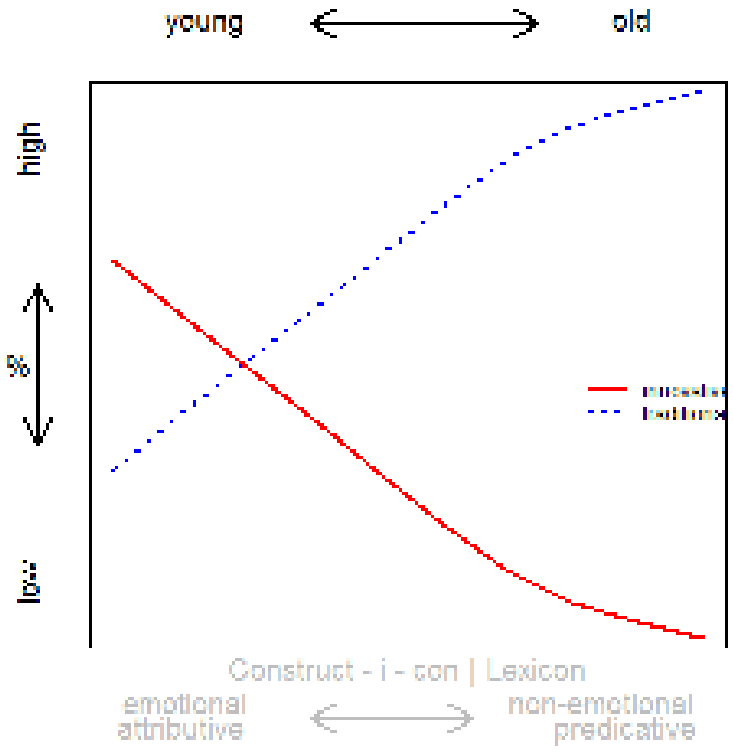
Diffusion of Innovations



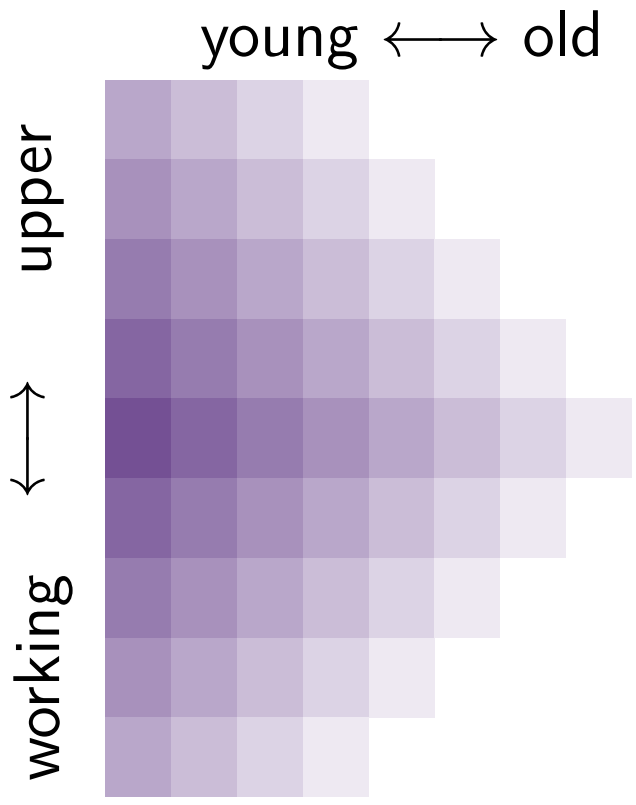


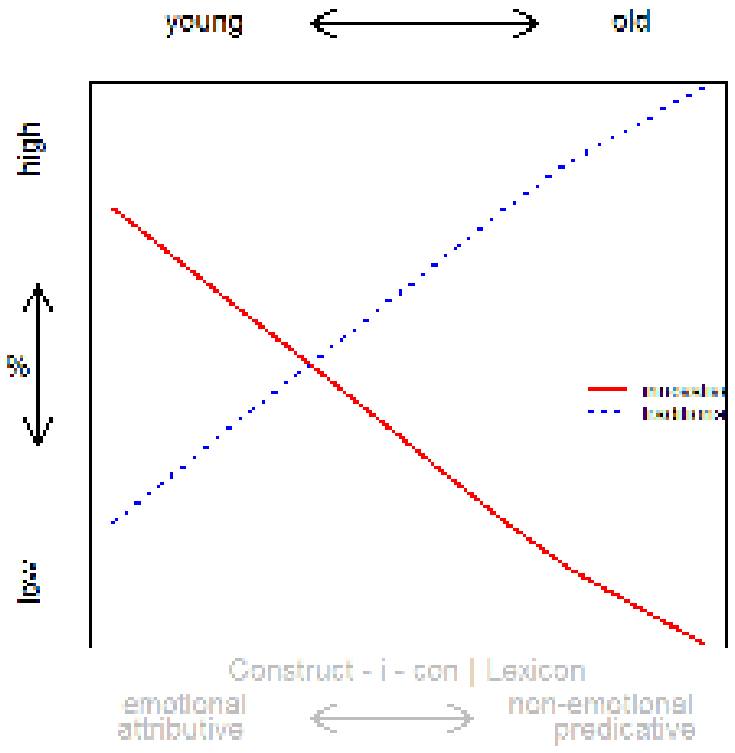
Diffusion of Innovations



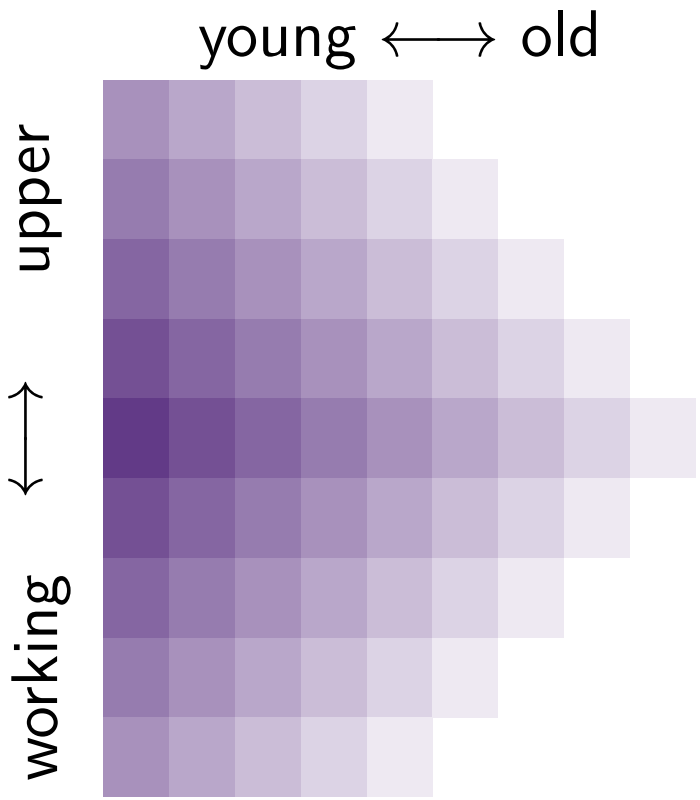


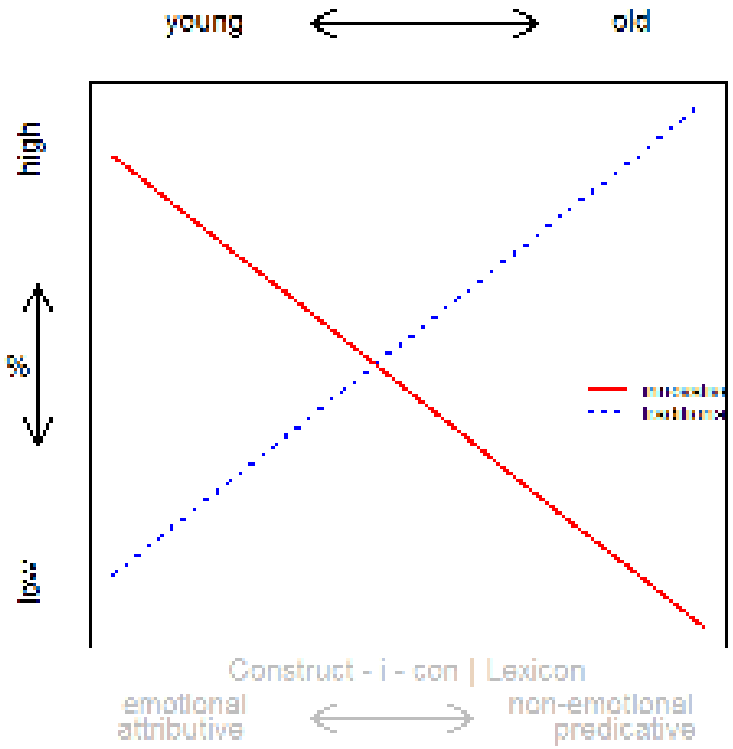
Diffusion of Innovations



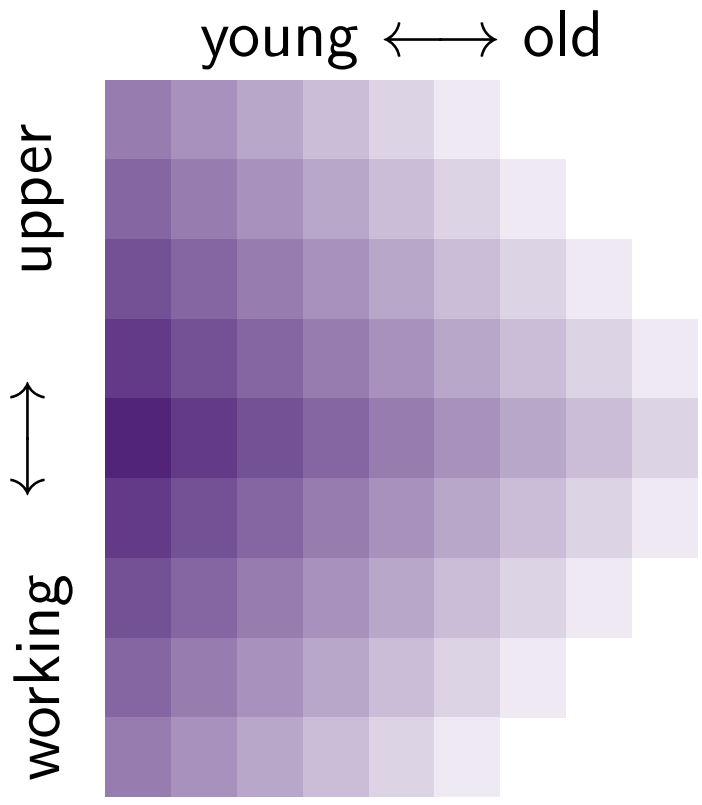


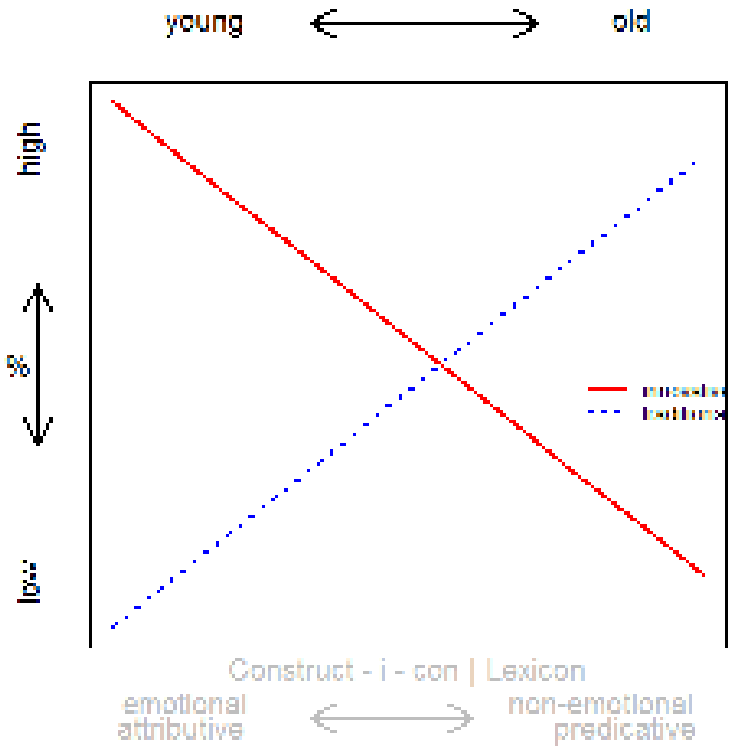
Diffusion of Innovations



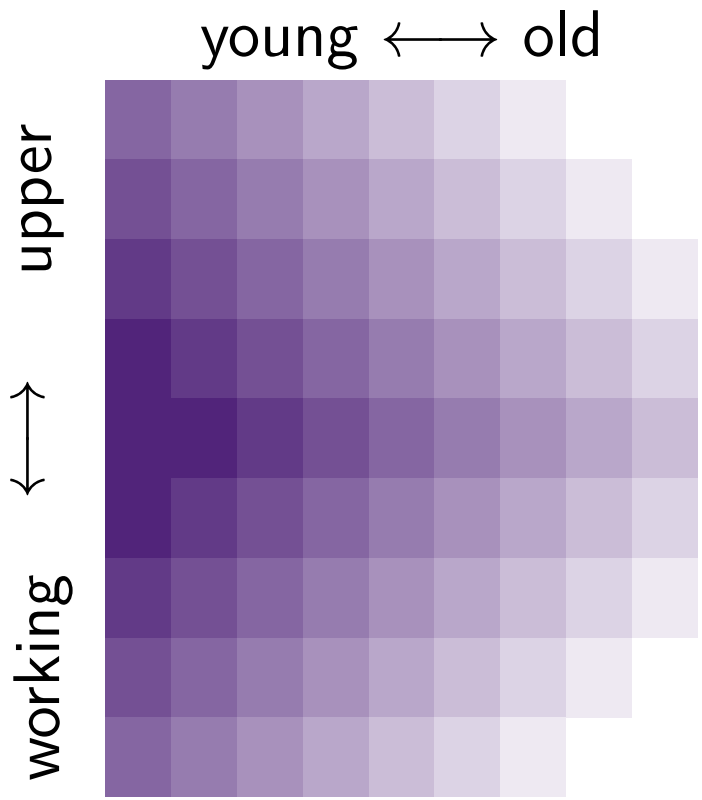


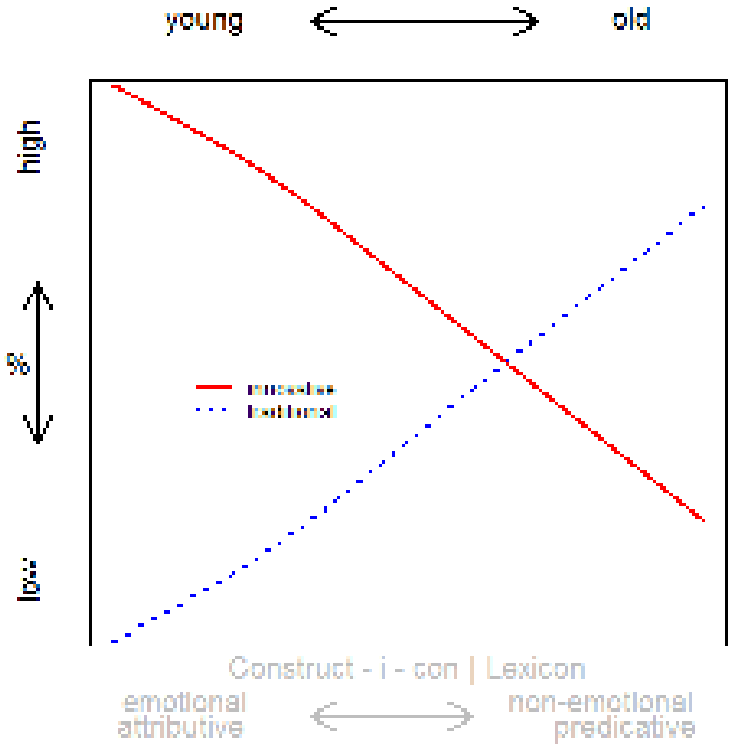
Diffusion of Innovations



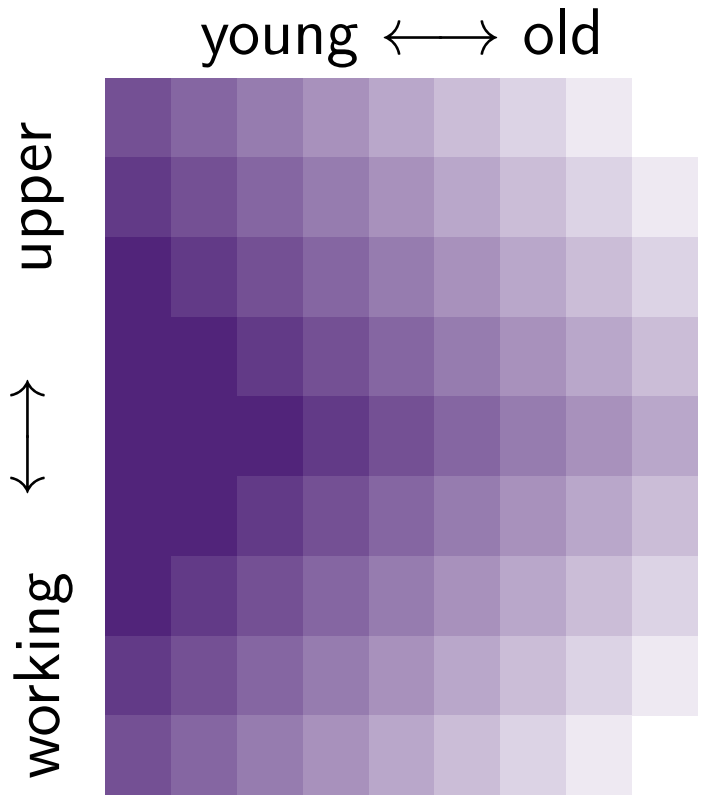


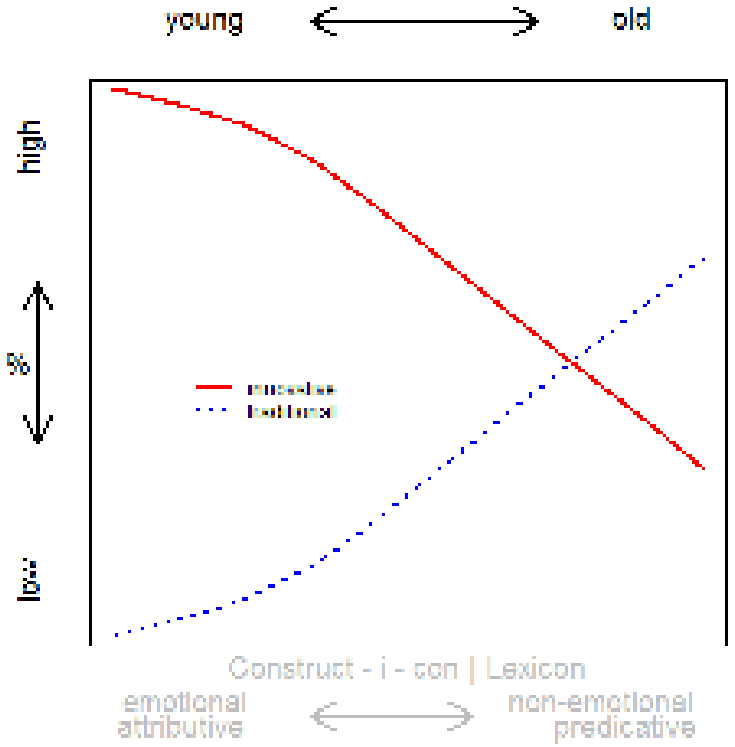
Diffusion of Innovations



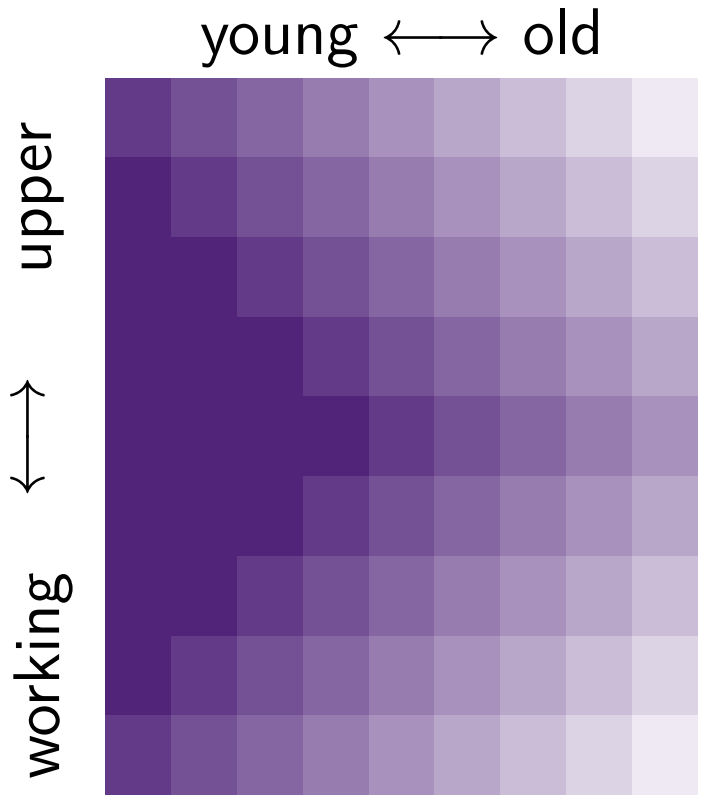


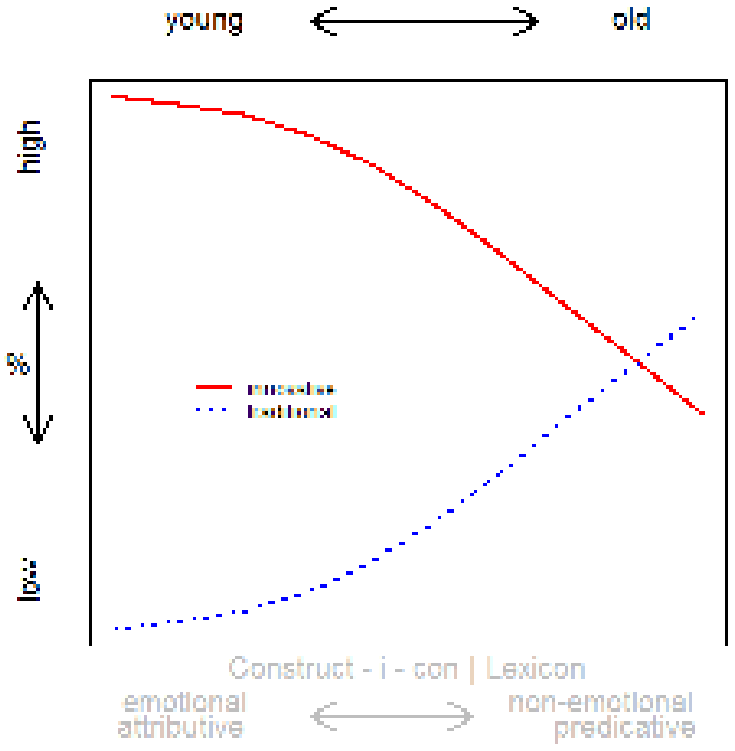
Diffusion of Innovations



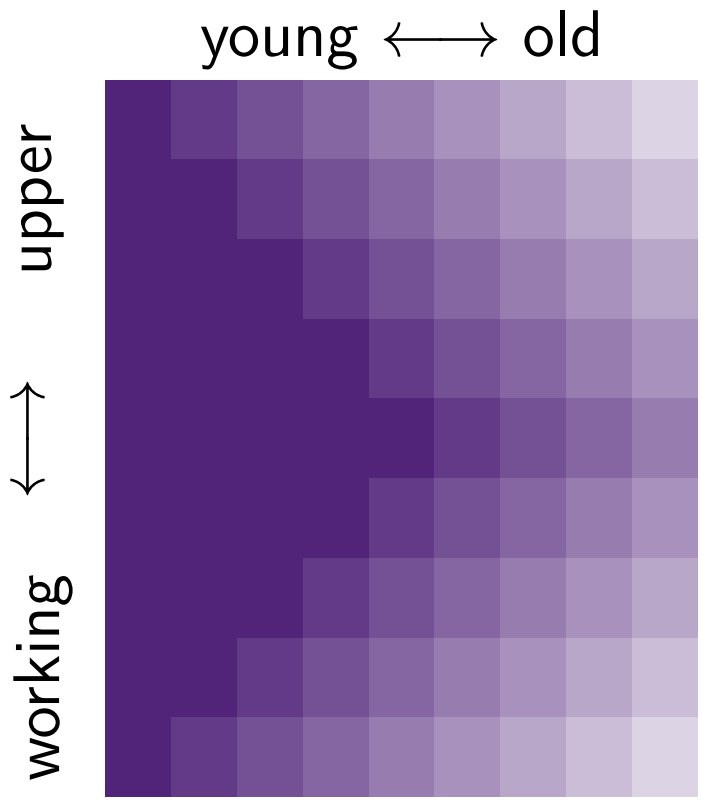


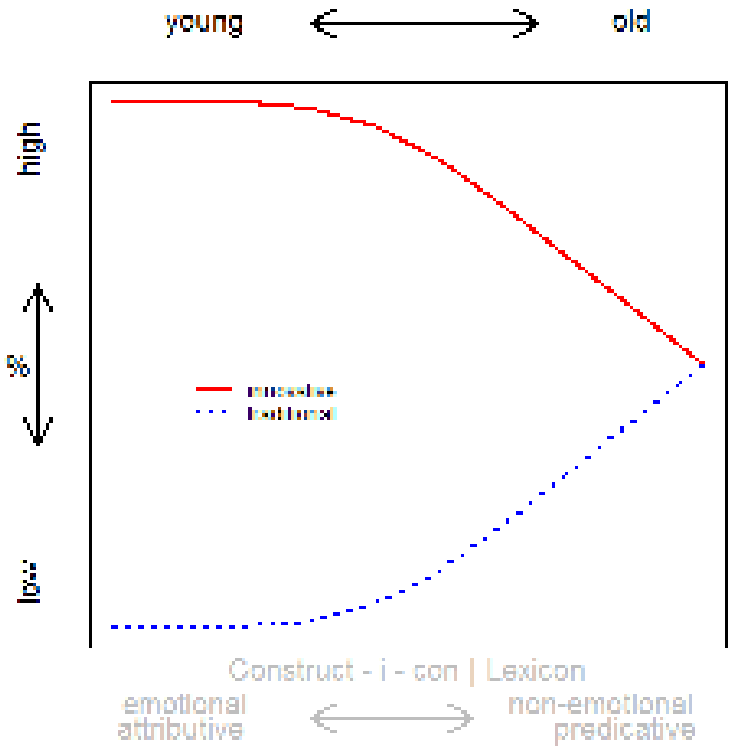
Diffusion of Innovations



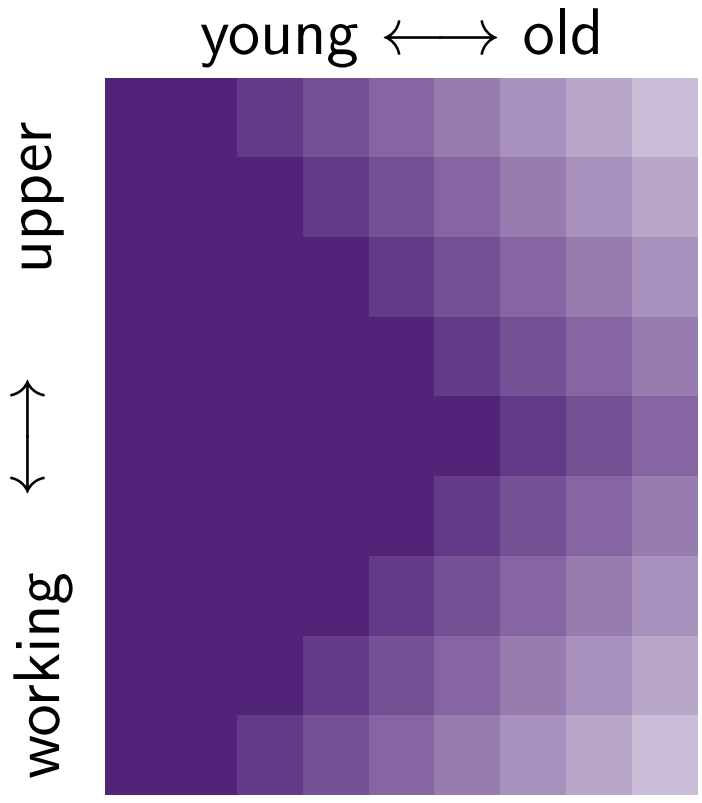


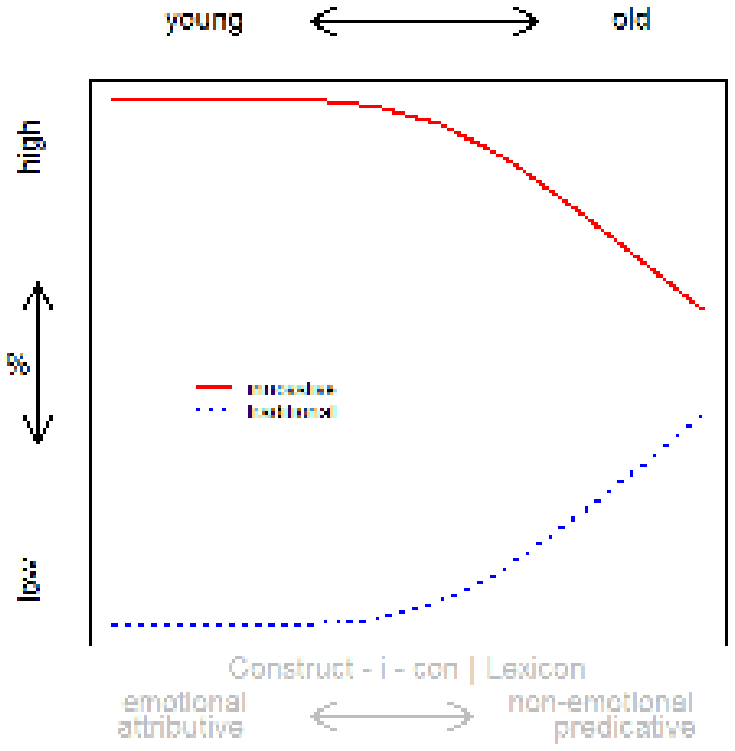
Diffusion of Innovations



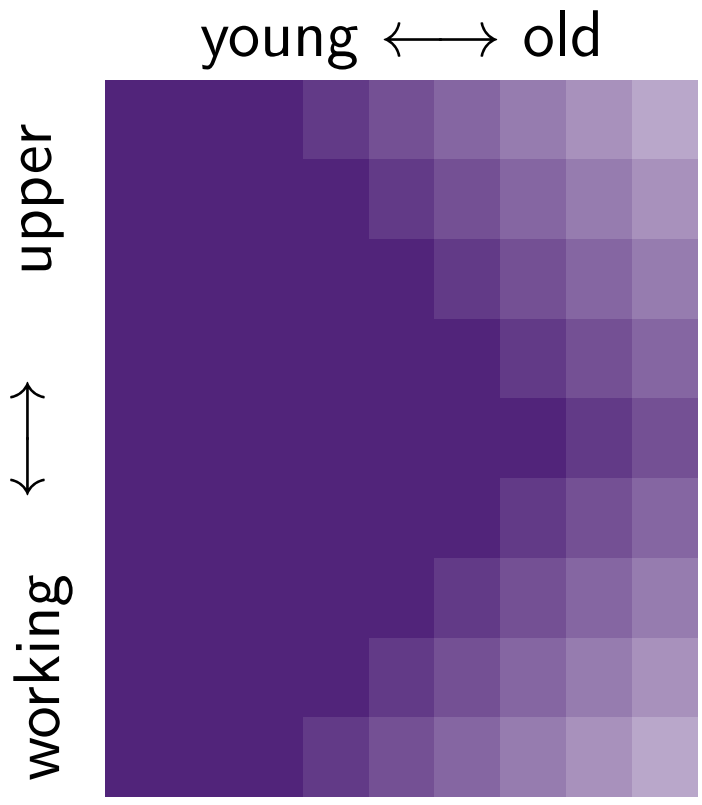


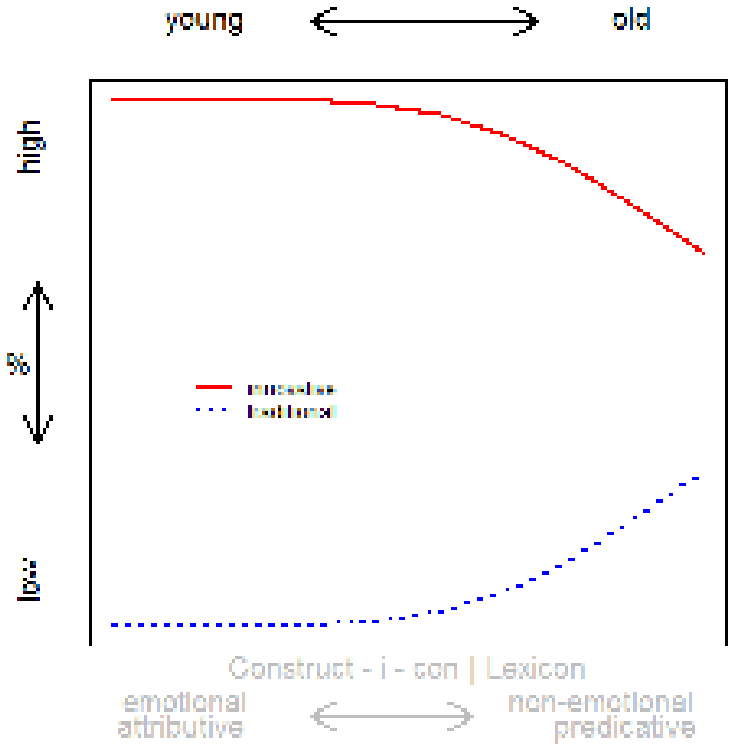
Diffusion of Innovations



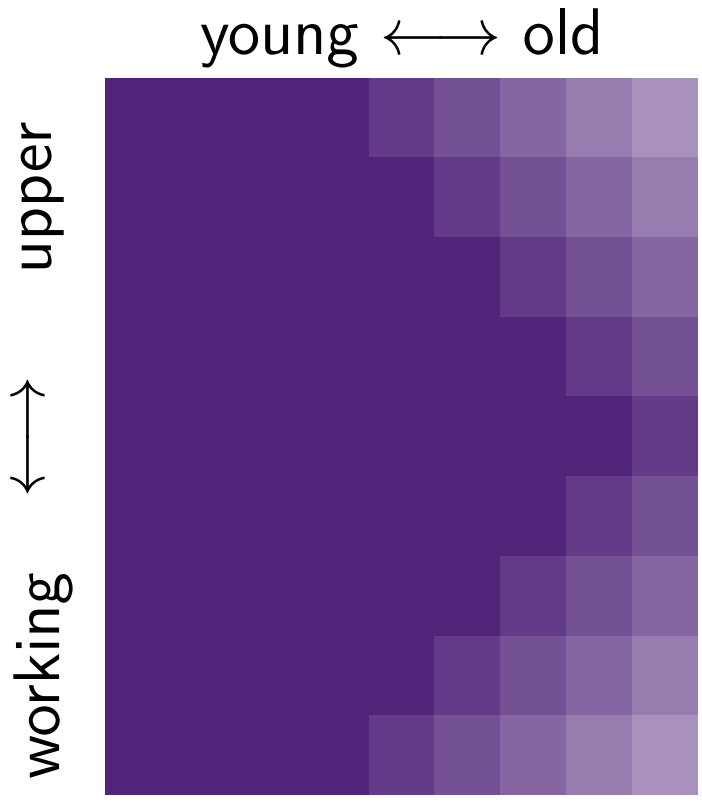


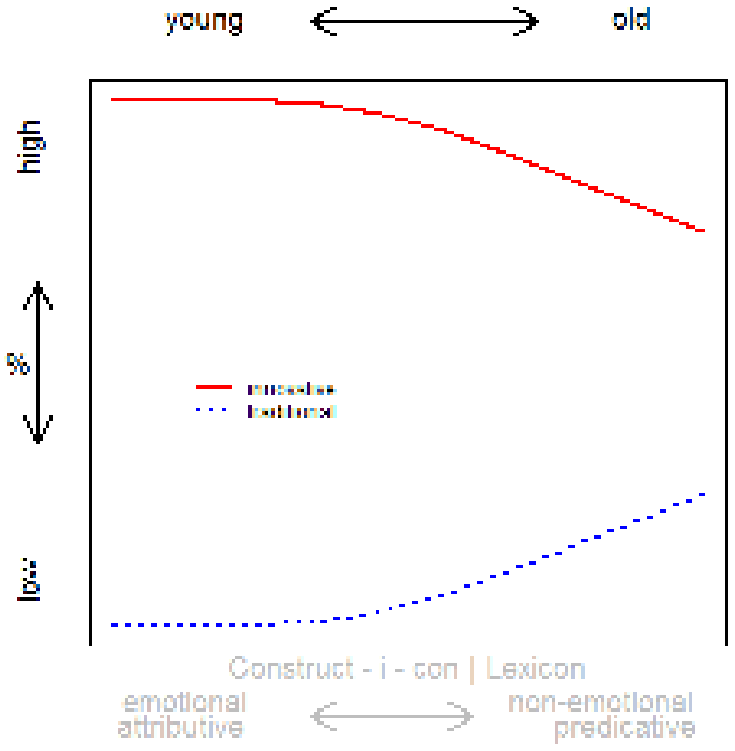
Diffusion of Innovations



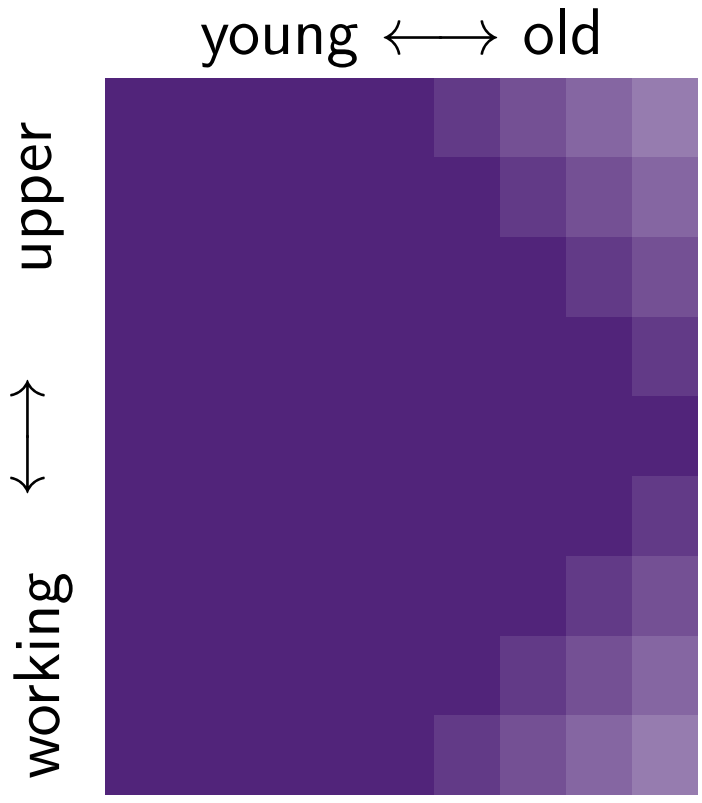


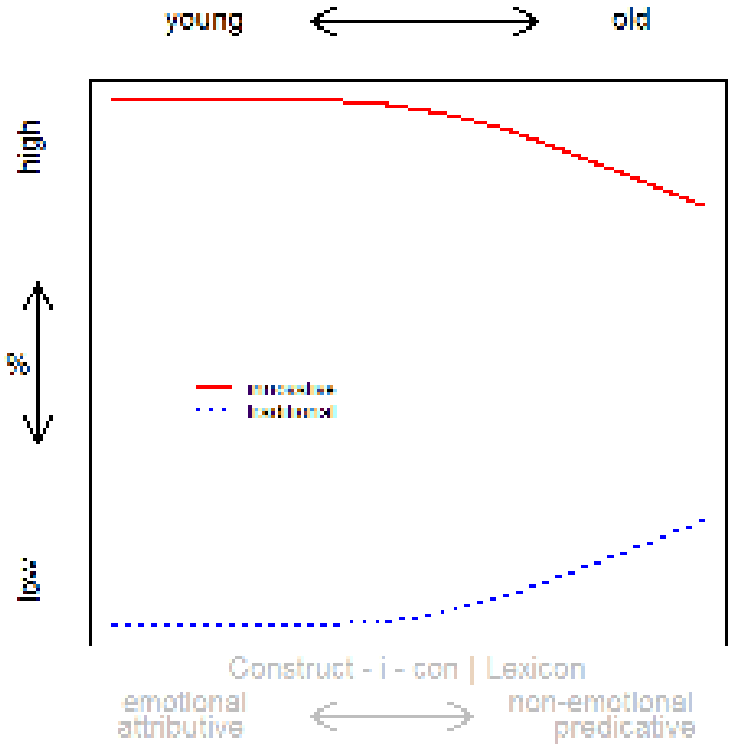
Diffusion of Innovations



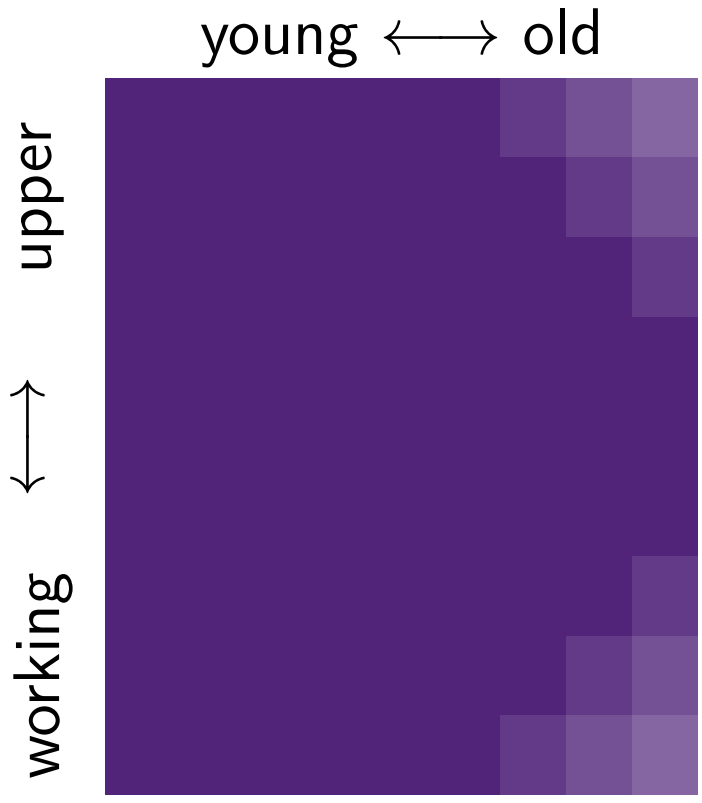


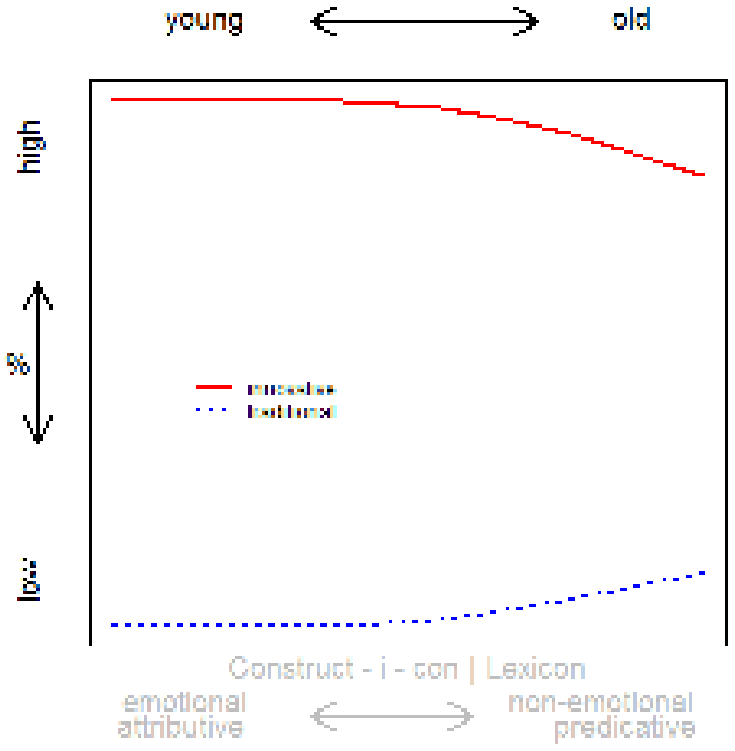
Diffusion of Innovations



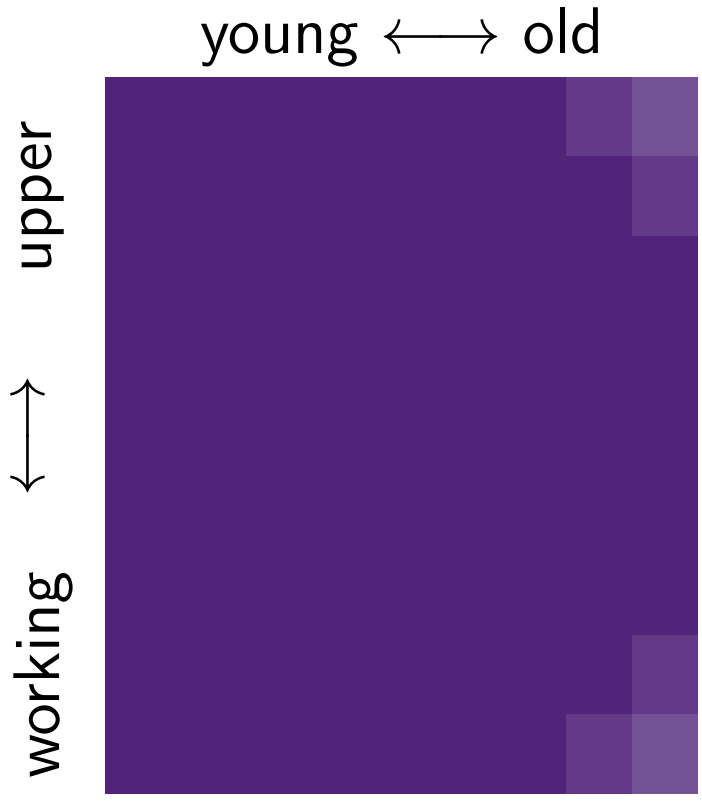


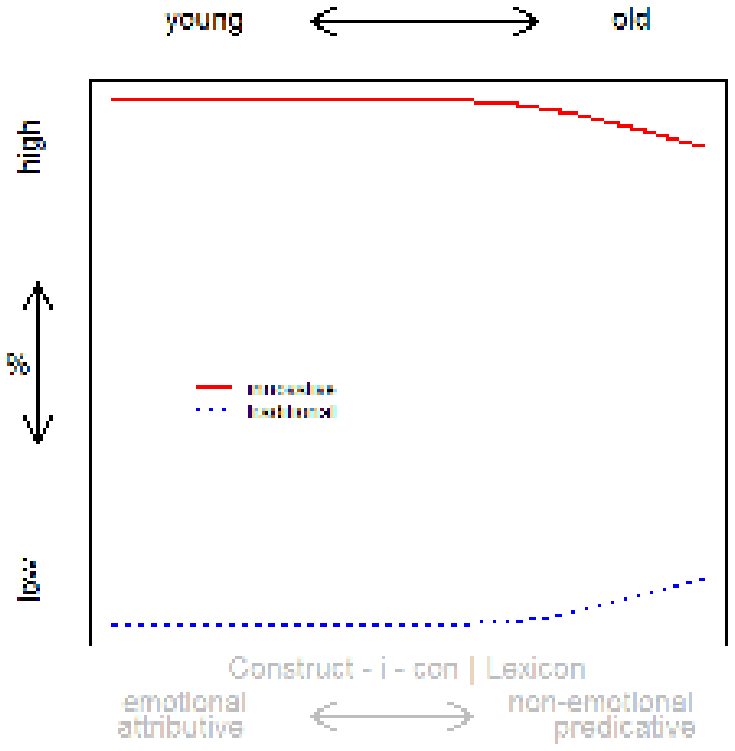
Diffusion of Innovations



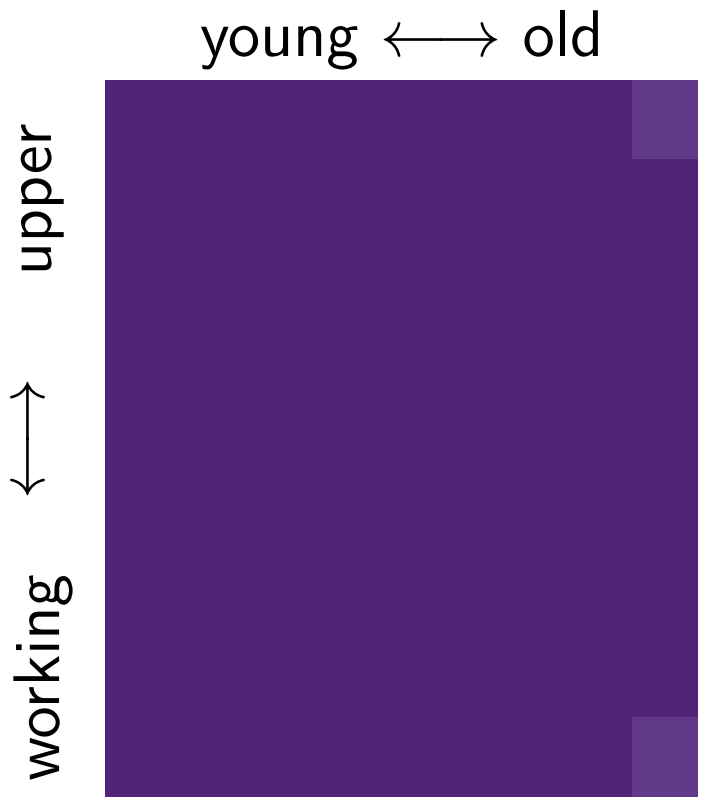


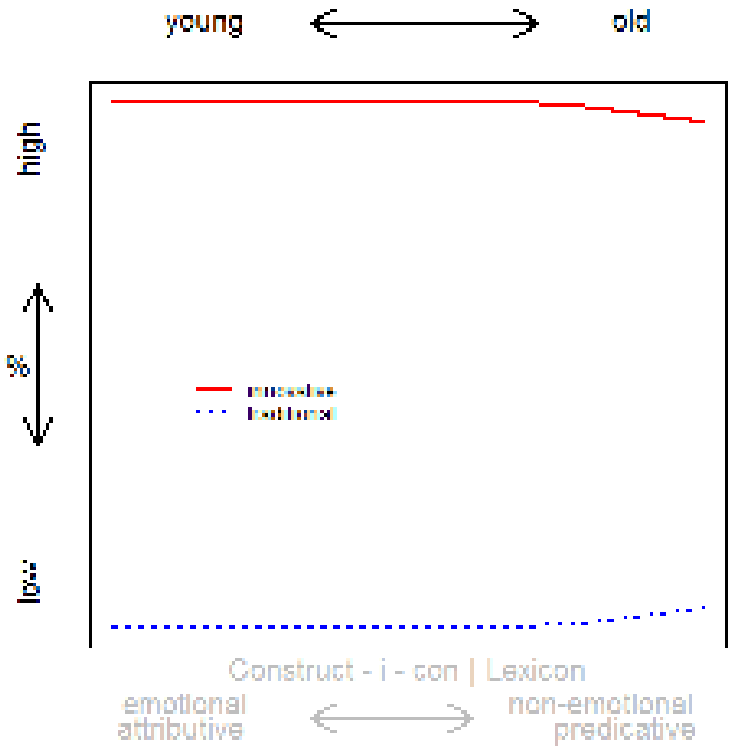
Diffusion of Innovations



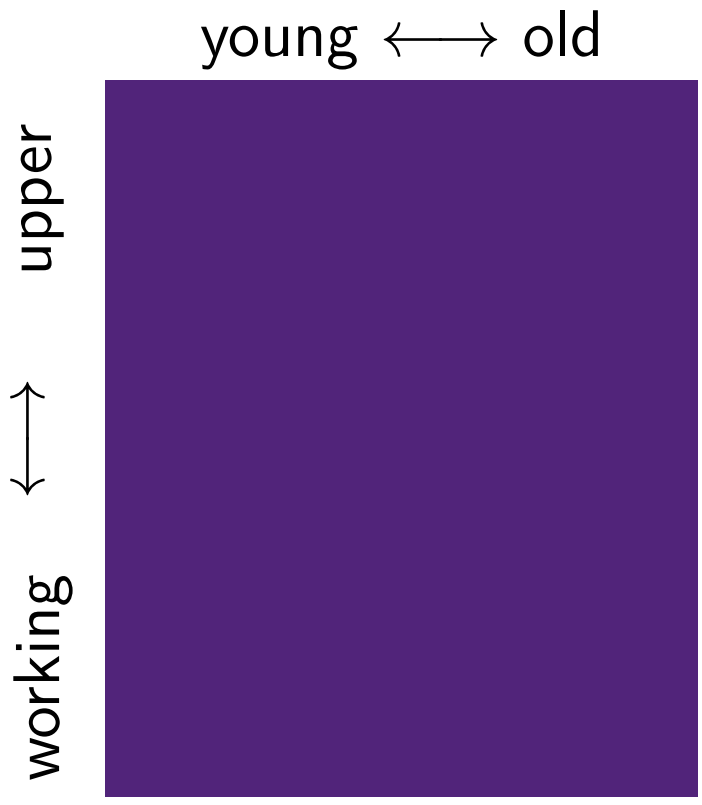


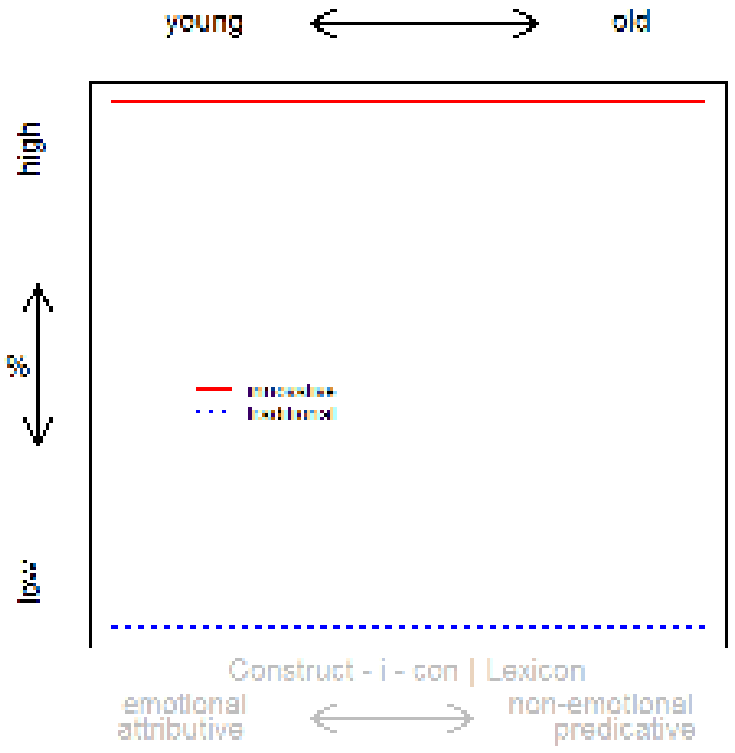
Diffusion of Innovations





Diffusion of Innovations





Q_6

Why is *very* replaced by *really* and not by any other variant (e.g. *so*, *quite*, *pretty*)?

→ What mechanisms underlie lexical replacement?

Scenario 1 (Broadening)

Really associate with many (but infrequent) adj. types

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

Argument

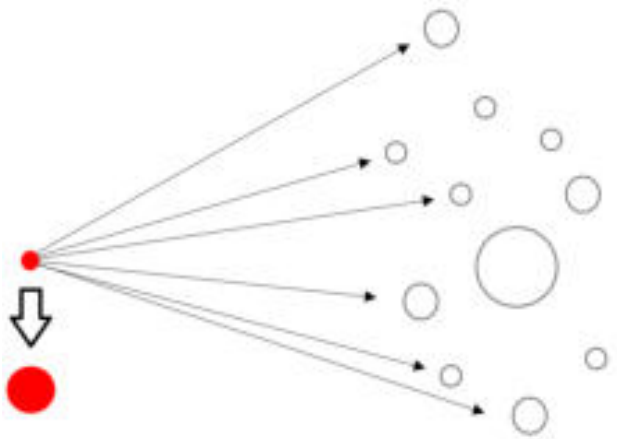
→ co-occurrence with many different adj. types

→ frequent use

→ deeper cognitive entrenchment

→ easier retrieval from memory

→ dominance within the amplifier system.



Prediction

Co-occurrence with many different adjective types

→ high lexical diversity

→ weak coll. attraction with specific adj. types

Scenario 2 (Specialization)

Really associate with few but frequent adj. types (HFAs)

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

Argument

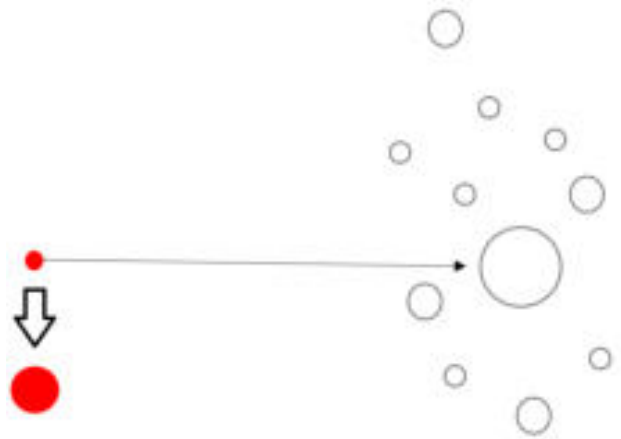
→ co-occurrence with high-freq. adj. types

→ frequent use

→ deeper cognitive entrenchment

→ easier retrieval from memory

→ dominance within the amplifier system.



Prediction

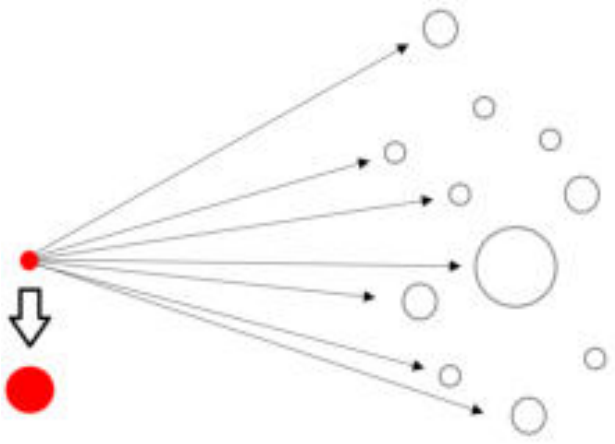
Co-occurrence with few high frequency adjectives

→ low lexical diversity

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

Scenario 3 (Randomness)

Really associate with random adj. types
→ We cannot predict which variants become successful based on their coll. profile.



H_1

If *really* is successful because of specialization on HFAs

$\rightarrow$ sig. pos. correlation with adjective frequency

If broadening $\rightarrow$ neg. correlation with adj. freq.
If random $\rightarrow$ no correlation with adj. freq.

Data and Methodology

International Corpus of English (ICE)

- Australian, British, Canadian, Irish, and New Zealand ICE components
- Shared design (allows meaningful comparisons between varieties of English)
- One million words (600,000 spoken and 400,000 written) from diverse spoken and written text types (cf. next slide) with each file containing app. 2,000 words.
- Accompanied by metadata and biodata of speaker (extremely interesting resource for variationist analyses)

Data Processing

- As described before (but only private dialogues)
- Sentiment Analysis of adjective types (Jockers 2017)
- Determined if the same amplifier type had occurred within a span of three adjective slots previously ($\rightarrow$ priming)
- Token freq. of adjective type by age group (Tagliamonte and Roberts 2005)
- Semantic classification of adjective (simplified version of Dixon (1977), cf. also D'Arcy (2015); Tagliamonte (2008))
- Manual cross–evaluation of automated classification
- Addition of demographic info. about speakers

Variable Coding

Dependent Variable(s)				
really	nominal	yes/no occurrence of pre-adjectival *really*		
Independent Variable(s)				
Age	ordinal	min. young | middle-aged | old	extra linguistic	
AudienceSize	nominal	Dyad | MultipleInterlocutors		
ConversationType	nominal	MixedSex | SameSex		
Gender	nominal	Female | Male		
(Education)	nominal	College | NoCollege		
Priming	nominal	prime | noprime		
Emotionality	categorical	negative | nonemotional | positive	intra linguistic	
Function	nominal	attributive | predicative		
SemanticCategory	categorical	semantic category of adj.		
Gradability	nominal	gradable | nongradable		
Adjective	categorical	bad | funny | good | interesting | nice | other		
Frequency	numeric	Frequency of adj. by age group		

Statistical Analysis

Mixed-Effects Binomial Logistic Regression

(Baayen 2008; Faraway 2016)

- Standard models for multivariate analyses
- Can handle nested/grouped data structure
- Easy multicollinearity detection
- Problems of MEBLoR
  - Cannot handle small data sets (well)
  - Extremely high β-error rate (Johnson 2009)
    - if sig. effect: ✓
    - if no sig. effect: ???

Statistical Analysis

Boruta Analysis (Kursa et al. 2010)

- Alternative to regressions that can handle small data sets
- Variable selection procedure
- Extension/improvement of random forest analysis
- Hundreds of forests are grown → distribution of parameters rather than singe values (higher reliability)
- Problems of Boruta
  - Ignores multicollinearity(!)
  - Does not model nested/grouped data structure

Results

Results AusE: Observed, Boruta, and MEBLoR

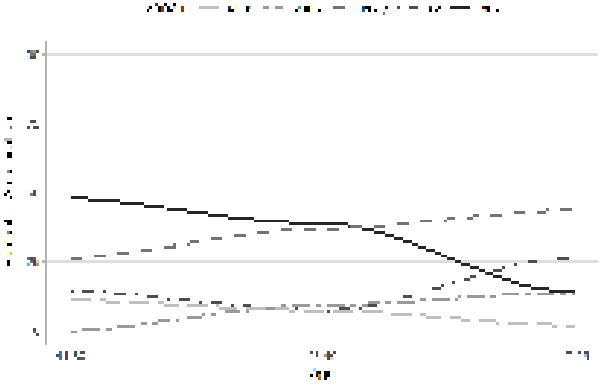


Figure 22: % Variants in AusE.

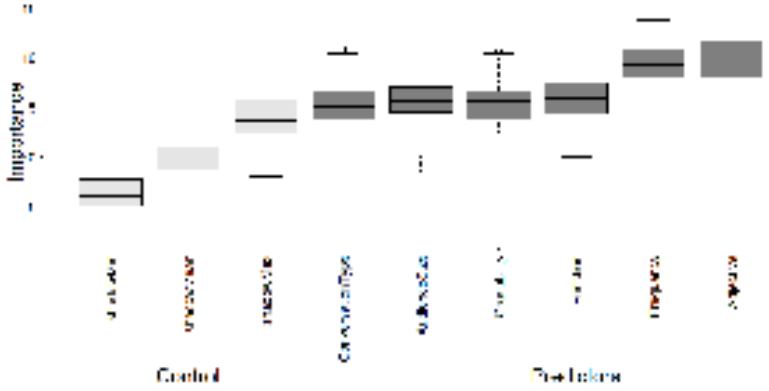


Figure 23: Boruta results for really in AusE.

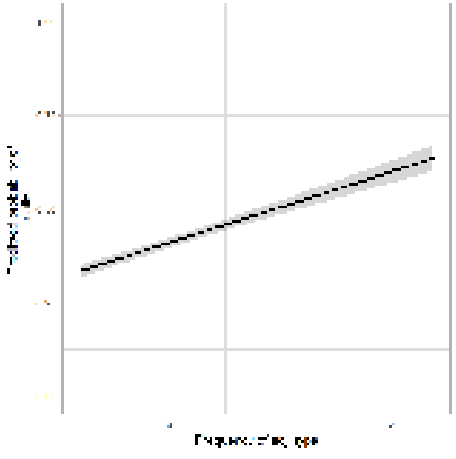


Figure 24: Prob. really in AusE by adj. freq.

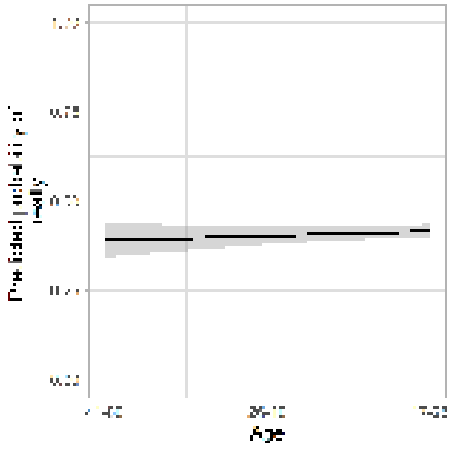


Figure 25: Prob. really in AusE across age.

Summary AusE Results

	Boruta		
Variety	**Age**	**Frequency**	**H_1?**
AusE	✗	✓	✓

Results BrE: Observed, Boruta, and MEBLoR

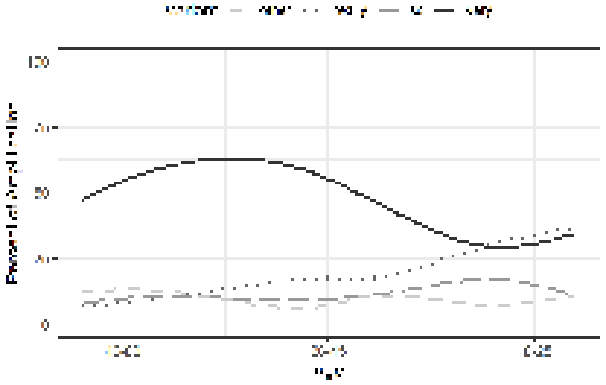


Figure 26: % Variants in BrE.

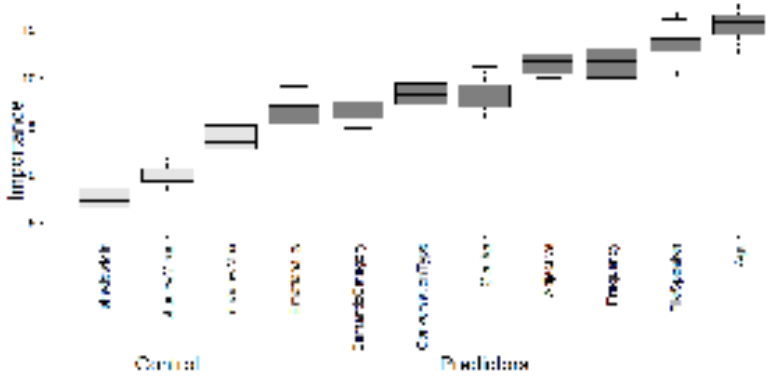


Figure 27: Boruta results for really in BrE.

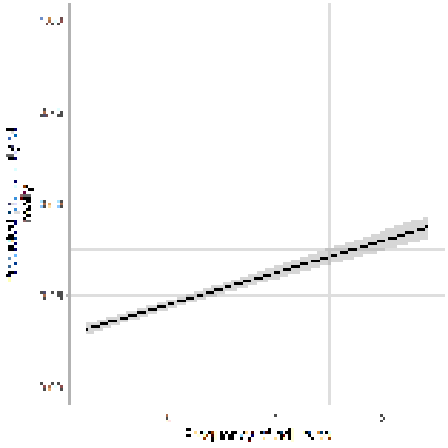


Figure 28: Prob. really in BrE by adj. freq.

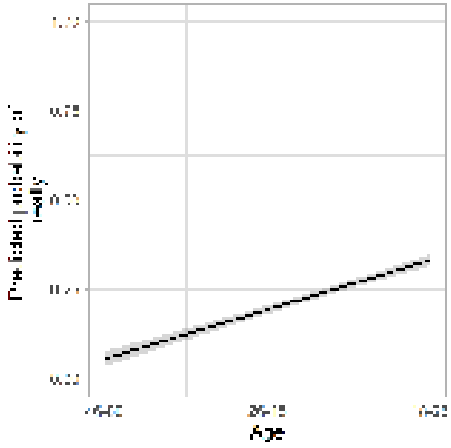


Figure 29: Prob. really in BrE across age.

Summary BrE Results

	Boruta		
Variety	**Age**	**Frequency**	**H_1?**
AusE	✗	✓	✓
BrE	✓	✓	✓

Results IrE: Observed, Boruta, and MEBLoR

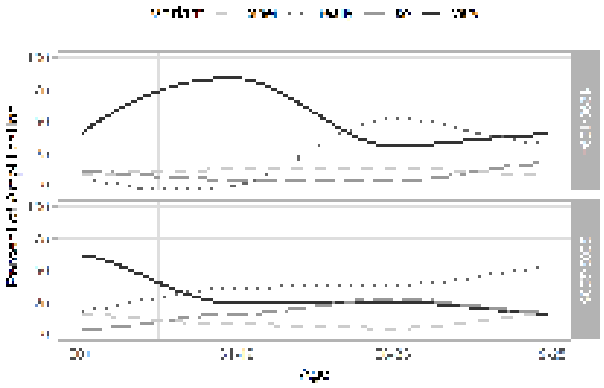


Figure 30: % Variants in IrE.

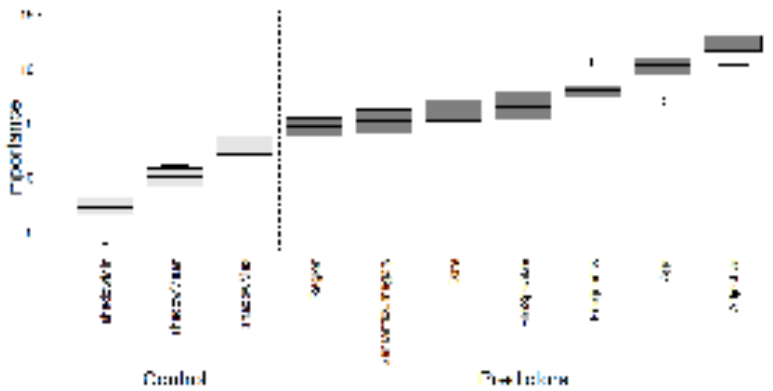


Figure 31: Boruta results for really in IrE.

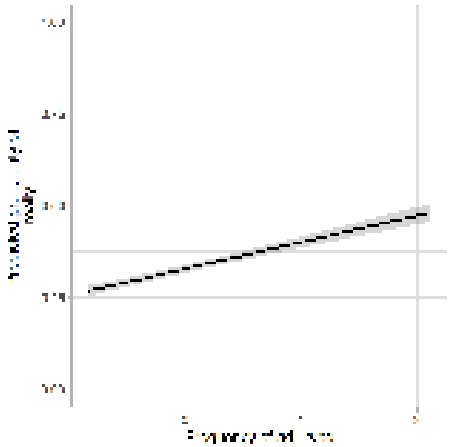


Figure 32: Prob. really in IrE by adj. freq.

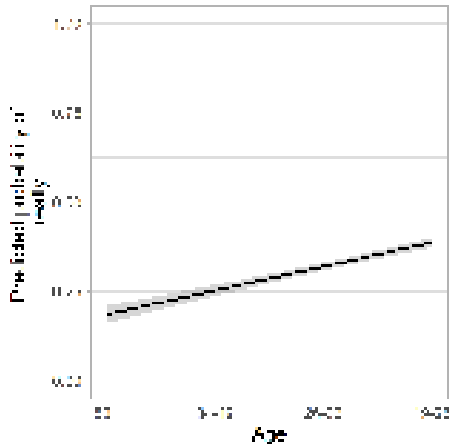


Figure 33: Prob. really in IrE across age.

Summary IrE Results

	Boruta		
Variety	**Age**	**Frequency**	**H_1?**
AusE	✗	✓	✓
BrE	✓	✓	✓
IrE	✓	✓	✓

Results NZE: Observed, Boruta, and MEBLoR

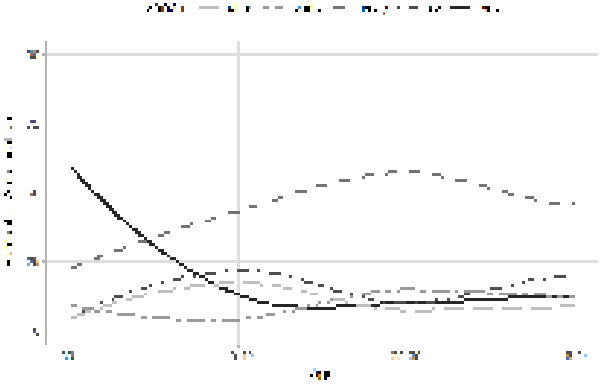


Figure 34: % Variants in NZE.

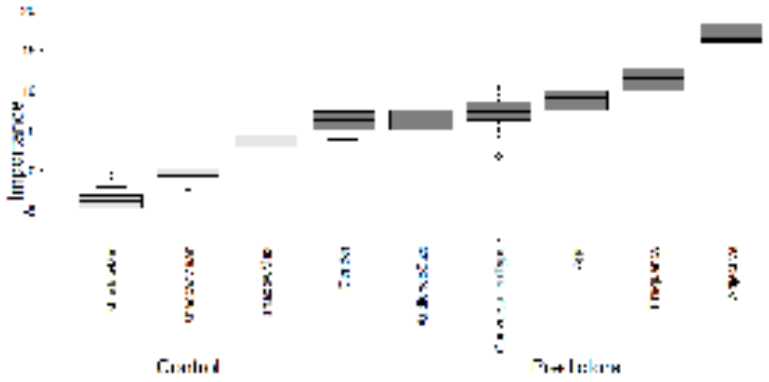


Figure 35: Boruta results for really in NZE.

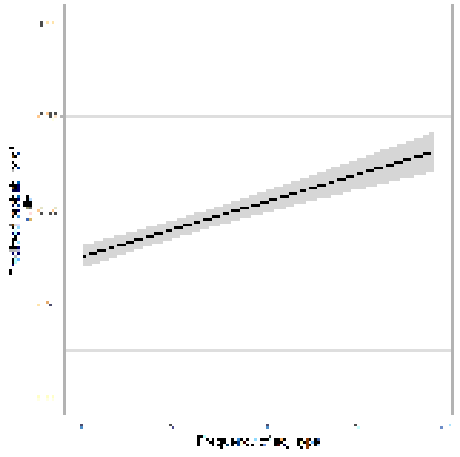


Figure 36: Prob. really in NZE by adj. freq.

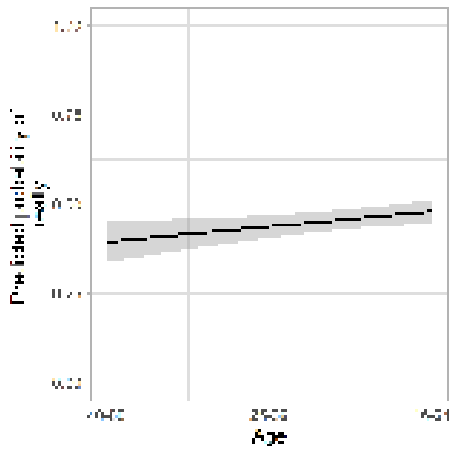


Figure 37: Prob. really in NZE across age.

Summary NZE Results

Variety	Boruta		H_1?
	Age	Frequency	
AusE	✗	✓	✓
BrE	✓	✓	✓
IrE	✓	✓	✓
NZE	✓	✓	✓

Results CanE: Observed, Boruta, and MEBLoR

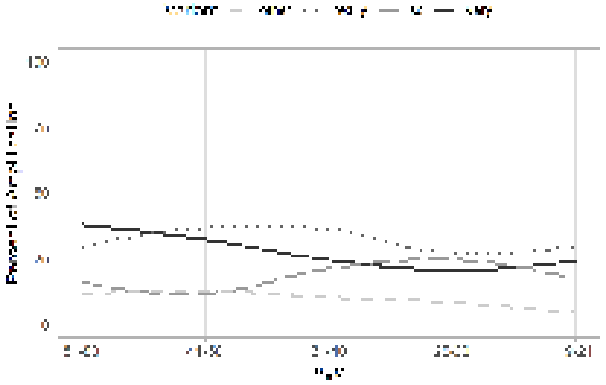


Figure 38: % Variants in CanE.

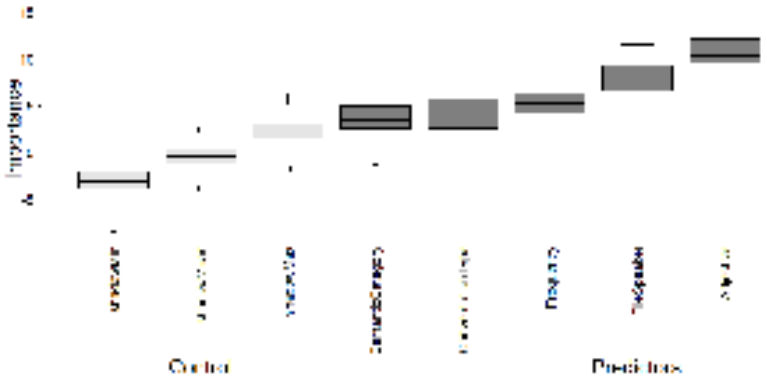


Figure 39: Boruta results for really in CanE.

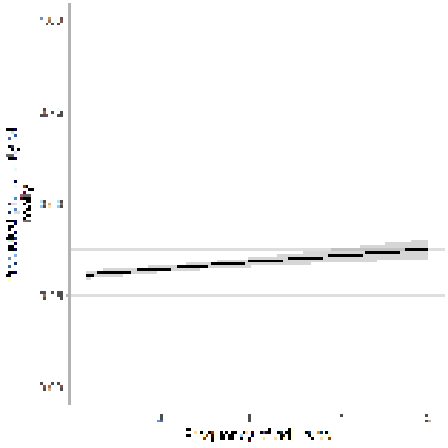


Figure 40: Prob. really in CanE by adj. freq.

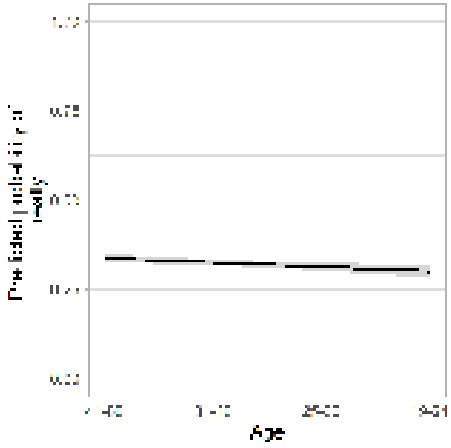


Figure 41: Prob. really in CanE across age.

Summary CanE Results

	Boruta		
Variety	**Age**	**Frequency**	**H_1?**
AusE	✗	✓	✓
BrE	✓	✓	✓
IrE	✓	✓	✓
NZE	✓	✓	✓
CanE	✗	✓	✓

Summary Results

Variety	Boruta		H_1?
	Age	Frequency	
AusE	✗	✓	✓
BrE	✓	✓	✓
IrE	✓	✓	✓
NZE	✓	✓	✓
CanE	✗	✓	✓

Discussion & Outlook

Summary

Results . . .

- confirm that *really* correlates with adj. freq. (positive correlation between the use of *really* and adjective frequency)
- suggest that lexical replacement is accompanied by (functional) re-organization in addition to diffusion through the speech community (absence of age effects)
  (see D'Arcy 2015)
- show that complementing mixed-modeling with Boruta is useful to avoid overlooking significant effects (avoidance of β-errors)

Discussion

- *Really* successfully replaced the dominant form *very* because it collocated with HFAs.
- No signs that *really* of broadening before taking over the system.
- Broadening once dominant (substantiates Tagliamonte and Denis 2014)

Argument

1. The co-occurrence with HFAs lead to the innovative variant being used as a more expressive variant to amplify certain HFAs.
2. The frequency of the innovative form increased because it piggybacked on the frequency of the HFA.
3. Increase in use $\rightarrow$ more deeply entrenched.
4. Deeper entrenchment $\rightarrow$ increased ease of retrieval.
5. Higher ease of retrieval $\rightarrow$ advantage over rival variants.
6. Innovative variant broadens because it increasingly co-occurs with more adj. types.

Discussion

Lexical replacement from a usage-based CxG perspective

- A default Cx loses its default status and is being replaced with a new default Cx.

Stage 1 Traditional stage (*very* dominant)	**Stage 2** Rivaling stage (no variant dominant)	**Stage 3** New stage (*really* dominant)
very good	really good	really good
very nice	very nice	really nice
very new	pretty new	really new
[very $X_{Adjective}$]	[$Y_{Intensifier}$] + [$X_{Adjective}$]	[really $X_{Adjective}$]

Interim outlook

Could this be a universal mechanism?

Test if the mechanisms. . .

- can be shown to have worked in analogous changes in English, e.g.
  3^{rd} p. sg. ind. morpheme: <eth> → <(e)s>
- can be shown to have worked in analogous changes in languages other than English

References

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.

Baayen, R. H. (2008). *Analyzing linguistic data. A practical introduction to statistics using R*. Cambridge: Cambridge University press.

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.

Dixon, R. M. W. (1977). Where have all the adjectives gone? *Studies in Language 1*, 19–80.

Faraway, J. J. (2016). *Extending the linear model with R: generalized linear, mixed effects and nonparametric regression models*, Volume 124. CRC Press.

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.

Gülzow, I. (2006). *The Acquisition of Intensifiers: Emphatic Reflexives in English and German Child Language*, Volume 22. Walter de Gruyter.

Hilpert, M. (2014). *Construction Grammar and its application to English*. Edinburgh: Edinburgh University Press.

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.

Jockers, M. L. (2017). Syuzhet 1.0.4 now on cran. http://www.matthewjockers.net/2017/12/16/syuzhet-1-0-4/.

Johnson, D. E. (2009). Getting off the goldvarb standard: Introducing rbrul for mixed-effects variable rule analysis. *Language and linguistics compass 3*(1), 359–383.

Kursa, M. B., W. R. Rudnicki, et al. (2010). Feature selection with the boruta package. *J Stat Softw 36*(11), 1–13.

Labov, W. (1964). Stages in the acquisition of standard english. In R. Shuy (Ed.), *Social Dialects and Language Learning*, pp. 77–103. Champaign, IL: National Council of Teachers of English.

Labov, W. (1966). *The Social Stratification of English in New York City*. Washington DC: Center for Applied Linguistics.

Labov, W. (1972). *Sociolinguistic patterns*. Philadelphia, PA: University of Pennsylvania Press.

Lieven, E. (2006). Producing multiword utterances. In E. V. Clark and B. F. Kelly (Eds.), *Constructions in Acquisition*, pp. 83–110. Stanford: CSLI Publications.

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.

Nardy, A., J.-P. Chevrot, and S. Barbu (2013). The acquisition of sociolinguistic variation: looking back and thinking ahead. *Linguistics 51*(2), 255–284.

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.

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

Schweinberger, M. (2017). Using intensifier-adjective bi-grams to investigate mechanisms of change. Paper presented at ICAME38. Prague, 27/5/2017.

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.

Tomasello, M. (2003). *Constructing a language: A usage-based theory of language acquisition*. Cambrigde: Harvard University Press.

Wagner, S. (2017a). Amplifier-adjective 2-grams world-wide: focus on pretty. Paper presneted at ICAME 37. Charles University Prague, 27/5/2017.

Wagner, S. (2017b). Totally new and pretty awesome: Amplifier–adjective bigrams in glowbe. *Lingua 200*, 63–83.

Thank you so, really, very much!

Acknowledgements

I would like to thank...

all ICE teams(!), in particular, Pam Peters and Adam Smith
for providing me with a preliminary version of ICE-Aus
(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

Absolutely fantastic! Adjective amplification in English

Dr. Martin Schweinberger

slides available at

www.martinschweinberger.de

R code upon request

The University Of Queensland

Mixed-Effects Binomial Logistic Regression

(Baayen 2008; Faraway 2016)

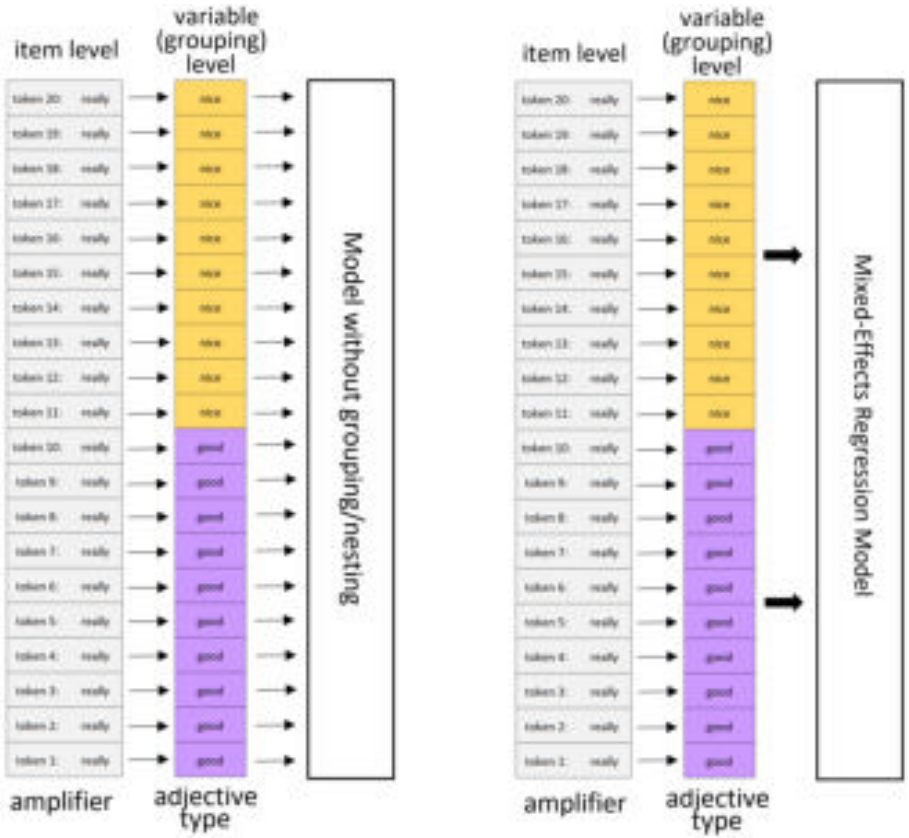


Figure 42: Difference between models without grouping/nesting and mixed-effects models (with grouping/nesting).