



**AALBORG UNIVERSITY**  
DENMARK

**Aalborg Universitet**

## **This construction is too hot to handle**

*A corpus study of an adjectival construction*

Jensen, Kim Ebensgaard

*Publication date:*  
2013

*Document Version*  
Accepted author manuscript, peer reviewed version

[Link to publication from Aalborg University](#)

*Citation for published version (APA):*

Jensen, K. E. (2013). *This construction is too hot to handle: A corpus study of an adjectival construction*. Paper presented at Annual Meeting of the Japanese Cognitive Linguistics Association, Kyoto, Japan.  
<http://homepage2.nifty.com/jcla/english/05.meeting.html>

### **General rights**

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal -

### **Take down policy**

If you believe that this document breaches copyright please contact us at [vbn@aub.aau.dk](mailto:vbn@aub.aau.dk) providing details, and we will remove access to the work immediately and investigate your claim.



# This construction is too hot to handle:

## A corpus study of an adjectival construction

Fourteenth Annual Meeting of the Japanese Cognitive Linguistics Association

Thematic session:

Corpus-based approaches to English adjectives: Meaning, grammar, and sound

Kim Ebensgaard Jensen  
CGS, Aalborg University



# Introduction

- The tatty furniture betrayed elegant lines, and the windows, *too **grimy** to see through*, stretched up ten feet. (COCA 2011 FIC Bk:NeverGentleman)
- They're *too **slow** to catch a seal in open water*. (COCA 2011 MAG NationalGeographic)
- The world has been *too **ready** to unlearn* the lessons of the financial crisis (COCA 2011 NEWS CSMonitor)
  - [*too* ADJ *to* V]-construction.



# Outline

- Scalar adjectival construction
- Data and method
- Qualitative analysis
- Collostructional analysis
  - Collexeme analysis
  - Covarying collexeme analysis
  - Distinctive collexeme analysis



# Scalar adjectival construction

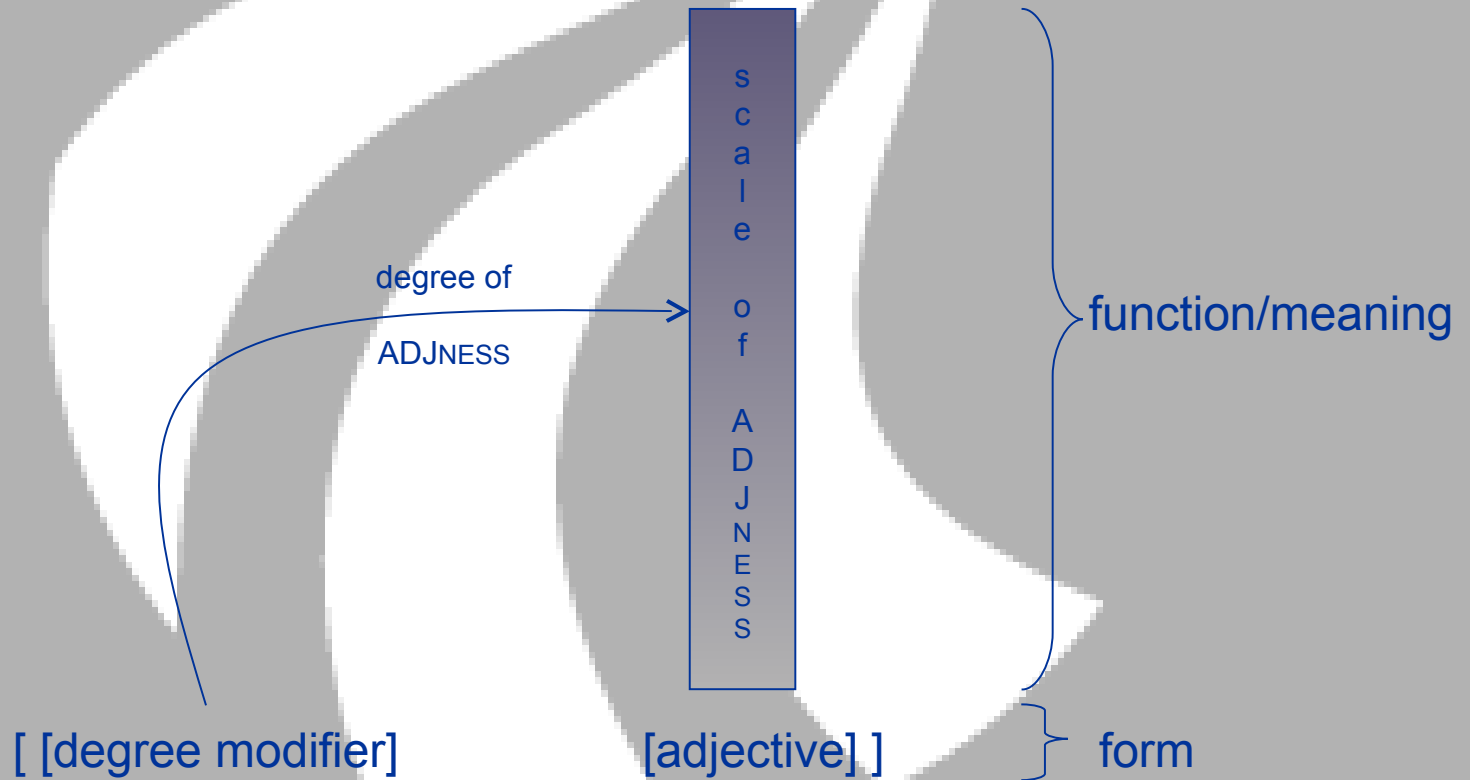
- Construction: “an entrenched routine ..., that is generally used in the speech community ... and involves a pairing of form and meaning” (Croft 2005: 274)
- Adjectival construction: construction in which an adjectival element plays a semantically/functionally pivotal part.
- Scalar adjectival construction: adjectival construction which draws on the scalarity of gradable, or gradably construed, adjectives.

# Scalar adjectival construction

- Components of a scalar adjectival construction:
  - Adjective slot: provides ADJNESS; draws on the content domain (Paradis 2000: 148)
  - Degree modifier: construes ADJNESS as a scale and specifies a degree of ADJNESS; draws on the schematic domain (Paradis 2000: 148-149)

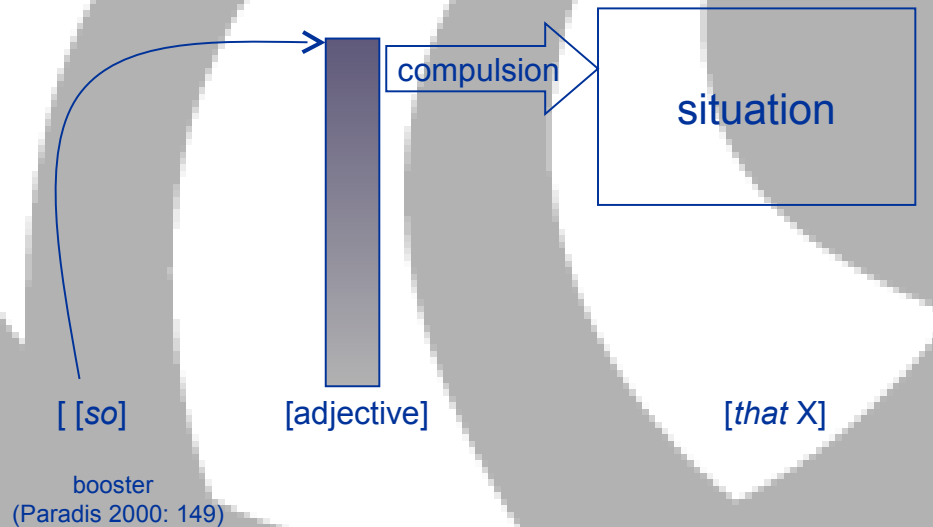
# Scalar adjectival construction

(Very) schematic structure



# Scalar adjectival construction

- An example: [so ADJ *that* X] (Bergen & Binsted 2003)
  - The film's ending was *so shocking that it physically hurt you.*
  - It was *so cold* in the kitchen *that there was frost on the lettuce.*





# Data and method

- Data:
  - 2011 segment of *Corpus of Contemporary American English (COCA)* (Davies 2013)
    - 20,445,868 words
    - Covers: fiction, magazines, newspapers, academic texts, speech
  - 1189 instances of [too ADJ to V] retrieved
- Method(s):
  - Qualitative analysis (identification of possible subcategories of use)
  - Quantitative analyses
    - Frequency analysis of identified categories
    - Collostructional analysis (using Gries [2007])
      - Collexeme analysis of items in ADJ-slot (Stefanowitsch & Gries 2003)
      - Covarying collexeme analysis of ADJ- and V-slots (Stefanowitsch & Gries 2004b, 2005)
      - Distinctive collexeme analysis of potential subcategories of use (Gries & Stefanowitsch 2004a).
    - Cluster analysis (using Jensen [2013])
      - Distribution of Dixon's (2004) adjective categories (drawing on an enhanced version of his typology)
      - Distribution of prototypically scalar/gradable and absolute/non-gradable items in the ADJ-slot
      - Using Canberra distancing and McQuitty clustering



# Data and method

- Why use COCA 2011 (or any corpus)?
  - Usage-based linguistics:
    - Language system is emergent (Hopper 1987, Kemmer & Barlow 2000)
    - Naturally occurring language reflects language system (Bybee & Hopper 2001, Tummerts et al. 2003)
  - Corpus linguistics has a number of scientifically attractive points (e.g. falsifiability, verifiability, objectivity, completeness etc. – see Kirk [1996: 253-254] for more, also Biber et al. [1998])



# Data and method

- Why use collocation analysis?
  - Principle of semantic compatibility: “words can (or are likely to) occur with a given construction if (or to the degree that) their meanings are compatible” (Stefanowitsch & Gries 2005: 4)
  - Principle of semantic coherence states that, “since a word in any slot of a construction must be compatible with the semantics provided by the construction for that slot, there should be an overall coherence among all slots” (Stefanowitsch & Gries 2005: 11).



# Qualitative analysis

- Syntactic form

co-text



co-text

degree modifier

infinitive clause

external properties

internal properties  
(Fillmore 1988: 36)

external properties



# Qualitative analysis

- Two category sets observed:
  - Two categories based on underlying semantic force-dynamic relations between the ADJ- and V-slots.
  - Three categories based on participant role selection via zero-anaphoric relations between infinitive clause and the immediate co-text of [*too* ADJ *to* V].

# Qualitative analysis: Force-dynamic categories

- Two force-dynamic categories:
  - Prevention: degree of  $ADJ_{NESS}$  prevents situation expressed by V-slot from occurring
  - Enablement: degree of  $ADJ_{NESS}$  enables situation expressed by V-slot to occur.

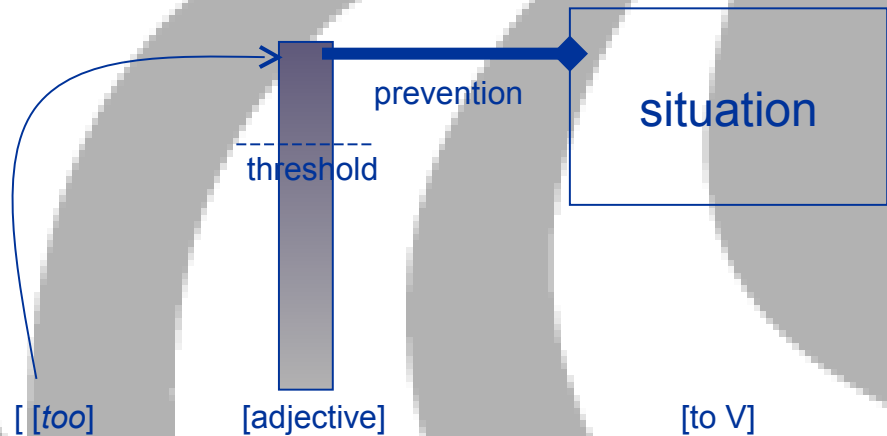
# Qualitative analysis: Force-dynamic categories

- Prevention:

- Most of them are *too young, too green to know just how human I am.* (COCA 2011 FIC Bk:FinalStorm)
- If the making of a revolution is drama, punctuated with tragedies *too numerous to count*, making peace is long-form prose requiring iterations of conversation between people. (COCA 2011 MAG TechReview)
- After all, when my children were preteenagers and *too young to handle last-minute flight cancellations or heavy turbulence on their own*, the programs offered considerable peace of mind. (COCA 2011 NEWS NYTimes)

# Qualitative analysis: Force-dynamic categories

- Prevention:





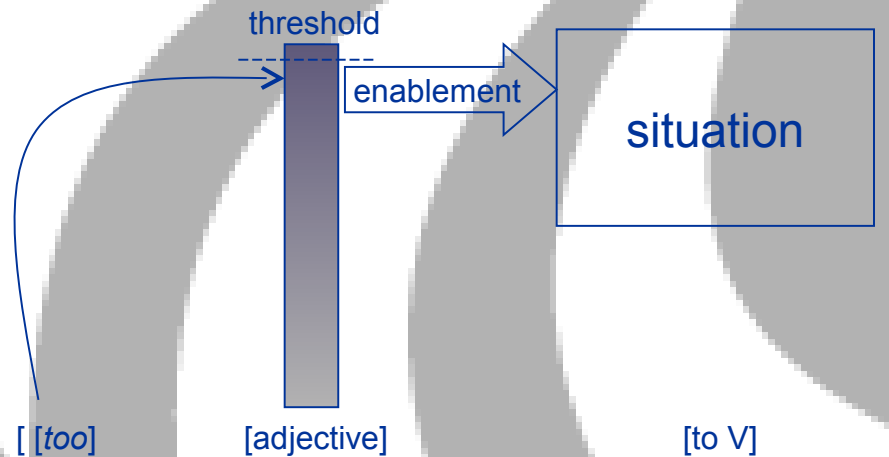
# Qualitative analysis: Force-dynamic categories

- Enablement:

- But at a time when our discourse has become so sharply polarized, at a time when we are far *too eager to lay the blame for all that ails the world at the feet of those who happen to think differently than we do.* (COCA 2011 SPOK CNN\_Situation)

# Qualitative analysis: Force-dynamic categories

- Enablement:



# Qualitative analysis: Force-dynamic categories

- Underlying force-dynamic cognitive models (Johnson 1987, Talmy 2000 413-470)
  - Prevention: BLOCKAGE image schema in which the ANTAGONIST (degree of ADJNESS) is force-dynamically stronger than the AGONIST (primary participant in proposition expressed by infinitive clause).
  - Enablement: ENABLEMENT image schema in which a force-input (degree of ADJNESS) strengthens the AGONIST (primary participant in proposition expressed by infinitive clause).

# Qualitative analysis: Participant role selection categories

- Three participant role selection categories:
  - Primary participant role: zero-anaphoric reference from unexpressed primary participant (AGENT, EXPERIENCER, EXISTENT, COGNIZER etc.) in scenario of infinitive clause to element in co-text of [*too* ADJ *to* V]
  - Secondary participant role: zero-anaphoric reference from unexpressed secondary participant (THEME, PATIENT, BENEFICIARY etc.) in scenario of infinitive clause to element in co-text of [*too* ADJ *to* V]
  - Condition/factor: zero-anaphoric reference from unexpressed CONDITION/FACTOR in scenario of infinitive clause to co-text of [*too* ADJ *to* V], including '*too* ADJ'.



# Qualitative analysis:

## Participant role selection categories

- Primary participant role
  - A \$25 donation to the IRC can supply one dehydrated child *who is too weak to eat or drink* with an IV kit and fluids for two days. (COCA 2011 MAG Redbook)
  - I'm too *depressed to see straight*. (COCA 2011 FIC RedCedarRev)
  - A presidential candidate who needs an image consultant to tell him it might not be a good idea to take a T-shirt-clad hottie on a yacht called Monkey Business is *too dumb to be president*, anyway. (COCA 2011 MAG Newsweek)

# Qualitative analysis: Participant role selection categories

- Primary participant roles

- ... supply one dehydrated child *who is too weak*  $\emptyset$  *to eat or drink*  $\emptyset$  with an IV kit ...

AGENT      ACT      THEME
- I'm too depressed*  $\emptyset$  *to see straight*  $\emptyset$  .

EXPERIENCER      EXPERIENCE      EXPERIENCED
- A presidential candidate *who needs ... is too dumb*  $\emptyset$  *to be president* ...

EXISTENT      RELATION      ATTRIBUTE

# Qualitative analysis: Participant role selection categories

- Secondary participant role
  - She had the smallest room, *too small to let*, a cupboard really, Patrick hadn't even noticed it was there when they first moved in. (COCA 2011 FIC SouthwestRev)
  - It's *too sophisticated to have been programmed by some punk teenager*. (COCA 2011 MAG PopMech)

# Qualitative analysis: Participant role selection categories

- Secondary participant roles

• She had the smallest room, *too small* Ø *to let* Ø , a cupboard really

AGENT ACT THEME

• It's too sophisticated Ø *to have been programmed* *by some punk teenager*.

THEME ACT AGENT



# Qualitative analysis: Participant role selection categories

- Condition/factor
  - It's *too dark to see her eyes ...* (COCA 2011 FIC BK:LimeCreekFiction)
  - Pa fell through the ice in March, but the ground was still *too frozen to dig a grave.* (COCA 2011 FIC BoysLife)
  - The data are *too noisy to chalk that trend up to increased rainfall.* (COCA 2011 MAG PopMech)

# Qualitative analysis: Participant role selection categories

- Condition/factor

- It's too dark    Ø    to see    her eyes    Ø  
EXPERIENCER    EXPERIENCE    EXPERIENCED    CONDITION/FACTOR
- ... but the ground was still too frozen    Ø    to dig    a grave    Ø  
AGENT    ACT    THEME    CONDITION/FACTOR
- The data are too noisy    Ø    to chalk    that trend ...    Ø  
AGENT    ACT    THEME    CONDITION/FACTOR

# Quantitative analysis: Overall category frequencies

Table 1: Force-dynamic relations\*

Force-dynamic relation	Frequency
Enablement	43
Prevention	1089

\* $p=3.38e-212$

Table 2: Participant role categories\*

Participant role category	Frequency
Primary	680
Secondary	262
Condition	190

\* $p=1.23e-11$  (data pooling used)

# Quantitative analysis: Collexeme analysis

Table 3: Top 30 attracted items in the ADJ-position (log-likelihood)

Rank	Lexeme	Collostruction strength	Rank	Lexeme	Collostruction strength
1	early	677.00759229121	16	afraid	99.6186614246447
2	busy	666.448099991938	17	frightened	92.1886545008188
3	young	623.149878585395	18	happy	89.2711341851548
4	late	574.118935192231	19	shy	83.0855196678116
5	big	346.708021280216	20	hot	79.0284195716512
6	good	317.591527127464	21	quick	66.0200941046014
7	old	279.666675859592	22	difficult	62.978771207314
8	weak	259.278759213788	23	dark	61.0548252891548
9	small	250.572920981746	24	preoccupied	60.9005457482469
10	tired	238.609461040527	25	nervous	58.2376111117956
11	scared	196.533754201181	26	short	54.8908679014025
12	expensive	135.097202239891	27	eager	53.3703002644430
13	drunk	119.118638140439	28	lazy	51.0469007261563
14	embarrassed	113.349312383912	29	dumb	50.7555562520501
15	heavy	111.934834777222	30	polite	50.4724975374736

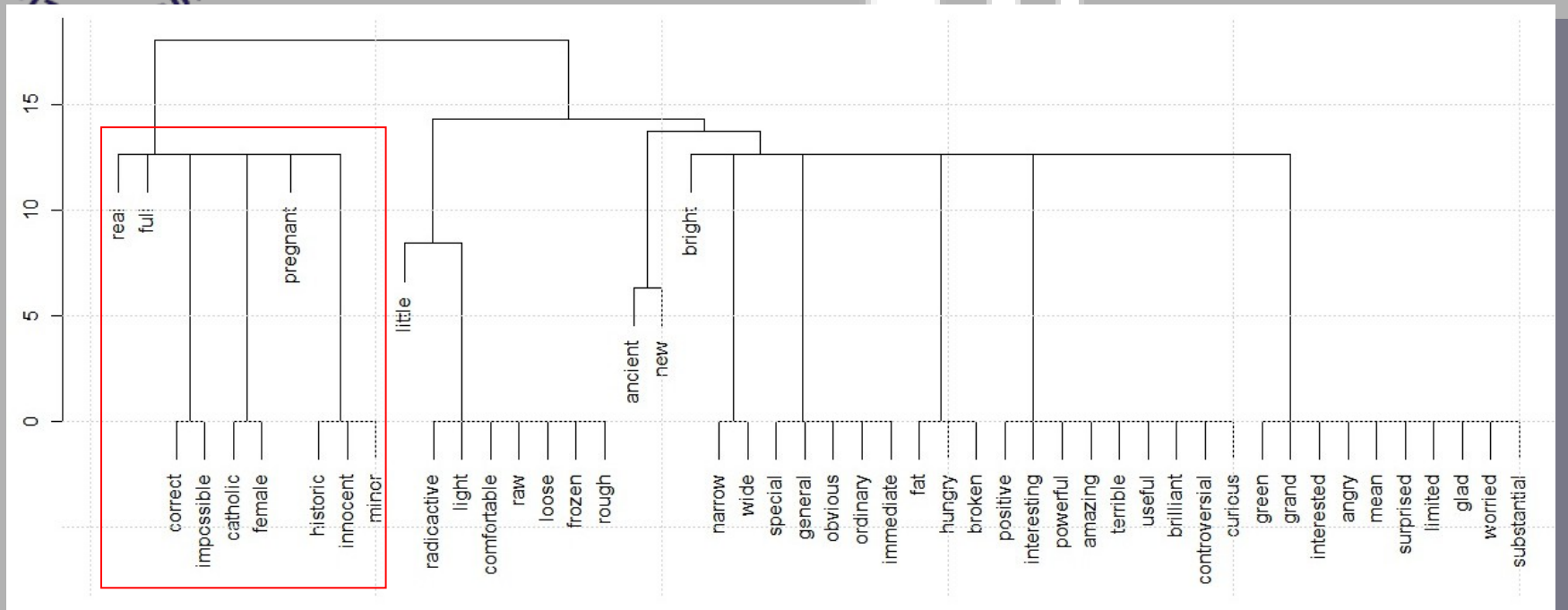
- Different semantic classes
- Scalarity as shared feature found at higher level of semantic categorization
- Technically an item-class-specific construction (Croft 2003: 57-58; Tomasello 2003: 139)

# Quantitative analysis: Collexeme analysis

Table 4: Bottom 50 attracted items in the ADJ-slot

Rank	Lexeme	Collostruction strength	Rank	Lexeme	Collostruction strength	Rank	Lexeme	Collostruction strength
1	frozen	5.3155567080975	18	glad	4.31249564578994	35	grand	3.51123470444561
2	controversial	5.29934983384863	19	narrow	4.27255459789068	36	amazing	3.50518111944987
3	curious	5.23197732942866	20	historic	4.23347020043966	37	bright	3.48864182634855
4	worried	5.16316496073382	21	useful	4.17087309431479	38	light	3.40395584432716
5	hungry	5.155671664543	22	limited	4.12744788841406	39	radioactive	3.08163286229869
6	brilliant	5.12231542095742	23	surprised	4.07876488372637	40	little	3.00189711473526
7	rough	5.11864533903977	24	Catholic	4.0313613348609	41	powerful	2.97770705107895
8	substantial	5.09315301750317	25	obvious	3.98715846428418	42	wide	2.97438934892024
9	innocent	5.00482856112027	26	mean	3.96545860175659	43	interesting	2.60615982600565
10	minor	4.93371714662497	27	terrible	3.94401837085054	44	positive	2.58953355804855
11	pregnant	4.8146437700432	28	ancient	3.9037870568314	45	green	2.16342315920390
12	ordinary	4.79911970141041	29	angry	3.84784271621698	46	general	2.16210806510439
13	loose	4.7745470362543	30	correct	3.81690436003222	47	special	1.71443894850748
14	immediate	4.67648130390031	31	comfortable	3.70166263059806	48	full	1.42785638901155
15	raw	4.61107266647797	32	impossible	3.61911360454218	49	real	0.840440686799456
16	broken	4.57236145453472	33	female	3.61750159129566	50	new	0.446494878623207
17	fat	4.44803971470921	34	interested	3.54028064688355			

# Quantitative analysis: Collexeme analysis



- Cluster of prototypically absolute/non-gradable/partially scalar adjectives (bottom 50 items in ADJ-position)
- Coerced (de Swart 2003) in [*too* ADJ *to* V]:
  - I am *too Catholic to be anything else*, but the church hierarchy tries my patience as nothing in my life ever has. (COCA 2011 MAC USCatholic)
  - Sunny was *too pregnant to argue*, but Jerry would have to sleep in the family room. (COCA 2011 FIC SouthwestRev)
  - I was *too female to be ruthless*. (COCA 2011 FIC Bk:DeadMansSwitch)

# Quantitative analysis: Covarying collexeme analysis

Table 5: Top 25 co-attracted lexeme pairs (log-likelihood)

Rank	ADJ	V	Collostruction strength
1	good	be	137.541973606126
2	big	fail	124.214583952871
3	early	tell	74.7141307967985
4	early	say	65.1619493422168
5	willing	compromise	32.0301027919704
6	precious	wear	30.3655661037372
7	late	change	30.1348488072069
8	numerous	count	28.7796928062779
9	happy	oblige	27.5455085556661
10	young	understand	25.2878140994993
11	dark	see	23.7372911654156
12	dangerous	release	22.6207239909575
13	heavy	lift	22.5634901460064
14	busy	bother	22.4968554430303
15	young	remember	22.4492007417720
16	early	gauge	22.2529058825033
17	quick	dismiss	21.8478093852388
18	excited	sleep	19.9901227360955
19	hot	sustain	18.5373716482582
20	disabled	stand	18.1884420908535
21	scared	talk	17.1332049639388
22	embarrassed	ask	16.5824396122544
23	eager	agree	16.1816901988028
24	acute	navigate	16.0625988650582
25	ancient	rouse	16.0625988650582

# Quantitative analysis: Covarying collexeme analysis

A number of pairs display obvious semantic coherence.

Table 5: Top 25 co-attracted lexeme pairs (log-likelihood)

Rank	ADJ	V	Collostruction strength
1	good	be	137.541973606126
2	big	fail	124.214583952871
3	early	tell	74.7141307967985
4	early	say	65.1619493422168
5	willing	compromise	32.0301027919704
6	precious	wear	30.3655661037372
7	late	change	30.1348488072069
8	numerous	count	28.7796928062779
9	happy	oblige	27.5455085556661
10	young	understand	25.2878140994993
11	dark	see	23.7372911654156
12	dangerous	release	22.6207239909575
13	heavy	lift	22.5634501480064
14	busy	bother	22.4968554430303
15	young	remember	22.4492007417720
16	early	gauge	22.2529058825033
17	quick	dismiss	21.8478093852388
18	excited	sleep	19.9901227360955
19	hot	sustain	18.5373716482582
20	disabled	stand	18.1884420908535
21	scared	talk	17.1332049639388
22	embarrassed	ask	16.5824396122544
23	eager	agree	16.1816901988028
24	acute	navigate	16.0625988650582
25	ancient	rouse	16.0625988650582



# Quantitative analysis: Covarying collexeme analysis

Table 5: Top 25 co-attracted lexeme pairs (log-likelihood)

Rank	ADJ	V	Collostruction strength
1	good	be	137.541973606126
2	big	fail	124.214583952871
3	early	tell	74.7141307967985
4	early	say	65.1619493422168
5	willing	compromise	32.0301027919704
6	precious	wear	30.3655661037372
7	late	change	30.1348488072069
8	numerous	count	28.7796928562779
9	happy	oblige	27.5455085556661
10	young	understand	25.2878140994993
11	dark	see	23.7372811654158
12	dangerous	release	22.6207239909575
13	heavy	lift	22.5634901480084
14	busy	bother	22.4968554430303
15	young	remember	22.4492007417720
16	early	gauge	22.2529058825033
17	quick	dismiss	21.8478093852388
18	excited	sleep	18.9901227360955
19	hot	sustain	18.5373716482582
20	disabled	stand	18.1884120908535
21	scared	talk	17.1332049639388
22	embarrassed	ask	16.5824396122544
23	eager	agree	16.1816901988028
24	acute	navigate	16.0625988650582
25	ancient	rouse	16.0625988650582

A number of pairs display obvious semantic coherence.

Others display semantic coherence that may be less obvious at first.

# Quantitative analysis: Covarying collexeme analysis

Table 5: Top 25 co-attracted lexeme pairs (log-likelihood)

Rank	ADJ	V	Collostruction strength
1	good	be	137.541973606126
2	big	fail	124.214583952871
3	early	tell	74.7141307967985
4	early	say	65.1619493422168
5	willing	compromise	32.0301027919704
6	precious	wear	30.3655661037372
7	late	change	30.1348488072069
8	numerous	count	28.7796928062779
9	happy	oblige	27.5455085556661
10	young	understand	25.2878140994993
11	dark	see	23.7372811654158
12	dangerous	release	22.6207239909575
13	heavy	lift	22.5634901480084
14	busy	bother	22.4968554430303
15	young	remember	22.4492007417720
16	early	gauge	22.2529058825033
17	quick	dismiss	21.8478093852388
18	excited	sleep	18.9901227360965
19	hot	sustain	18.5373716482582
20	disabled	stand	18.1884120908635
21	scared	talk	17.1332049639388
22	embarrassed	ask	16.5824396122544
23	eager	agree	16.1816901988028
24	acute	navigate	16.0625988650582
25	ancient	rouse	16.0625988650582

A number of pairs display obvious semantic coherence.

Others display semantic coherence that may be less obvious at first.

This indicates that there are underlying semantic relations between the ADJ- and V-positions.

# Quantitative analysis: Covarying collexeme analysis

Table 5: Top 25 co-attracted lexeme pairs (log-likelihood)

Rank	ADJ	V	Collostruction strength
1	good	be	137.541973606126
2	big	fail	124.214583952871
3	early	tell	74.7141307967985
4	early	say	65.1619493422168
5	willing	compromise	32.0301027919704
6	precious	wear	30.3655661037372
7	late	change	30.1348488072069
8	numerous	count	28.7796928062779
9	happy	oblige	27.5455085556661
10	young	understand	25.2878140994993
11	dark	see	23.7372811654158
12	dangerous	release	22.6207239909575
13	heavy	lift	22.5634501480084
14	busy	bother	22.4968554430303
15	young	remember	22.4492007417720
16	early	gauge	22.2529058825033
17	quick	dismiss	21.8478093852388
18	excited	sleep	18.9901227360965
19	hot	sustain	18.5373716482582
20	disabled	stand	18.1884120908635
21	scared	talk	17.1332049639388
22	embarrassed	ask	16.5824396122544
23	eager	agree	16.1816901988028
24	acute	navigate	16.0625988650582
25	ancient	rouse	16.0625988650582

A number of pairs display obvious semantic coherence.

Others display semantic coherence that may be less obvious at first.

This indicates that there are underlying semantic relations between the ADJ- and V-positions.

The force-dynamic relation categories are also reflected in many of the pairs.

# Quantitative analysis: Distinctive collexeme analysis

Table 6: Top 30 distinctive collexemes in enablement and prevention categories (log-likelihood)

Rank	Lexeme	Preferred category	Collostruction strength	Rank	Lexeme	Preferred category	Collostruction strength
1	happy	Enablement	89.260926217908	16	small	Prevention	2.83708788714208
2	willing	Enablement	46.9387441340836	17	weak	Prevention	1.88095345135307
3	eager	Enablement	40.0831186240193	18	tired	Prevention	1.80175341058188
4	quick	Enablement	37.5343608571090	19	scared	Prevention	1.4068437340666
5	anxious	Enablement	19.8243266312754	20	expensive	Prevention	1.17076555828738
6	ready	Enablement	19.8243266312754	21	heavy	Prevention	1.09221664490188
7	easy	Enablement	13.1695906448345	22	afraid	Prevention	0.935333748765715
8	likely	Enablement	13.1695906448345	23	hot	Prevention	0.935333748765715
9	early	Prevention	5.77098531119765	24	drunk	Prevention	0.856999504879296
10	young	Prevention	5.77098531119765	25	dark	Prevention	0.778736556949665
11	late	Prevention	4.61810055299727	26	difficult	Prevention	0.778736556949665
12	busy	Prevention	4.37305208570327	27	embarrassed	Prevention	0.778736556949665
13	good	Prevention	4.21007354855397	28	hard	Prevention	0.700544775298872
14	big	Prevention	3.88504026637943	29	large	Prevention	0.700544775298872
15	old	Prevention	3.15817662048135	30	short	Prevention	0.700544775298872

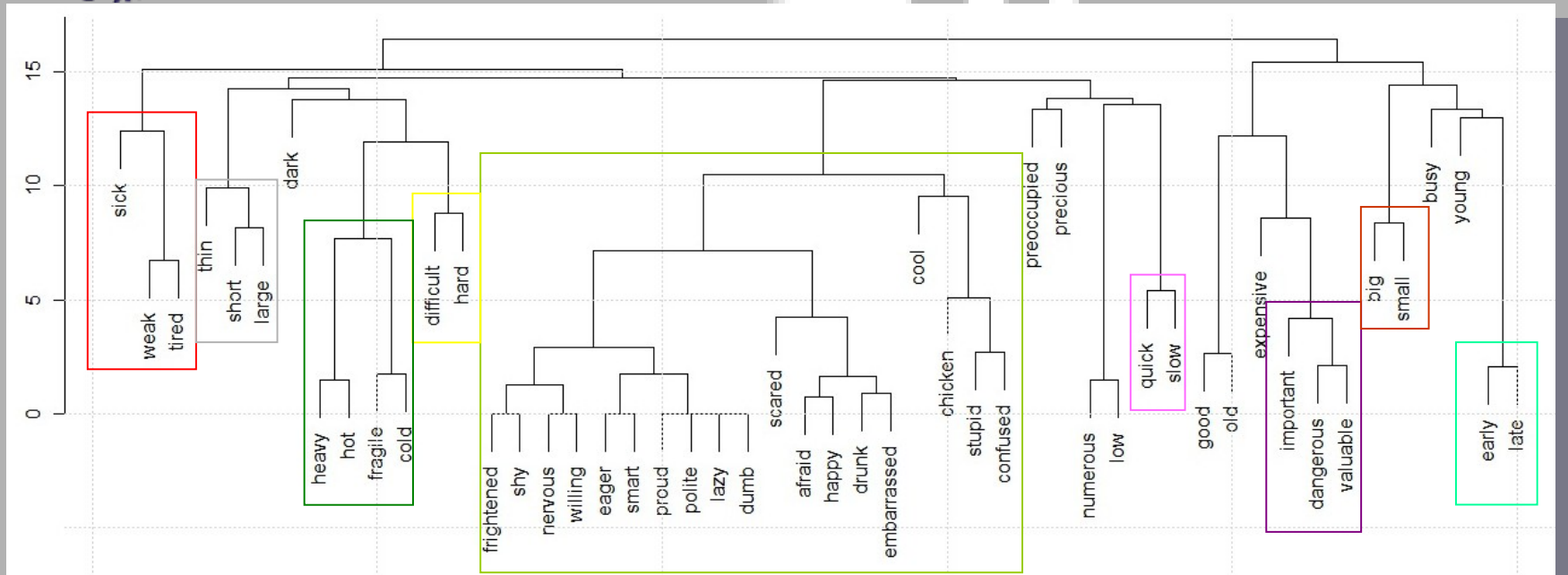
- Enablement category preferred by small number of adjectives
- Primarily human propensity adjectives (Dixon 2004)
- More strongly associated with enablement category than the adjectives that prefer the prevention construction
- Item-class-specific subconstruction? (Croft 2003: 57-58; Tomasello 2003: 139)

# Quantitative analysis: Distinctive collexeme analysis

Table 7: 50 distinctive collexemes in the participant role selection constructions (+ attraction, - repulsion)

Lexeme	Condition coll.strength	Primary coll.strength	Secondary coll.strength	Lexeme	Condition coll.strength	Primary coll.strength	Secondary coll.strength
early	53.2256635599884	-26.7089494378901	-8.20392136609225	embarrassed	-0.798729814981983	2.20953692219219	-1.13943352306837
late	40.3221007145987	-19.5735502922859	-6.60871443379653	quick	-0.718856833483785	1.98858322997297	-1.02549017076153
busy	-3.31004175396871	9.3131627933689	-5.02384632818973	costly	-0.239618944494595	-1.19785818914248	1.91046629276152
good	-4.23326801940451	5.3327957977823	-1.60997830558955	painful	-0.239618944494595	-1.19785818914248	1.91046629276152
difficult	-0.798729814981983	-3.99286063047495	6.36822097587174	frightened	-0.638983851985587	1.76762953775375	-0.911546818454695
young	-3.67260531305997	5.26026932530789	-1.73545155394760	shy	-0.638983851985587	1.76762953775375	-0.911546818454695
dark	6.04357240763952	-3.99286063047495	-0.537373531740405	deep	0.283320986977477	-1.59714425218998	1.39094118948877
hard	-0.718856833483785	-3.59357456742745	5.73139887828457	nervous	-0.559110870487389	1.54667584553453	-0.797603466147858
tired	-1.83707857445856	5.08193492104204	-2.62069710305725	willing	-0.559110870487389	1.54667584553453	-0.797603466147858
hot	0.892718038362304	-4.79143275656994	2.79655876563028	dim	1.54941800794525	-0.79857212609499	-0.227886704613674
old	-3.19491925992793	3.81538654919967	-1.12811447155400	mid	1.54941800794525	-0.79857212609499	-0.227886704613674
scared	-1.43771366696757	3.97716645994594	-2.05098034152306	eager	-0.479237888989191	1.32572215331531	-0.68366011384102
numerous	-0.479237888989191	-2.39571637828497	3.82093258552305	great	0.372577870301867	-1.19785818914248	0.869073607603298
heavy	-1.11822174097478	-1.76718974693157	3.77525120497288	complicated	-0.159745962996397	-0.79857212609499	1.27364419517435
weak	-1.15009953443425	3.16198850473023	-1.82082660298036	cute	-0.159745962996397	-0.79857212609499	1.27364419517435
cold	3.18089234156717	-2.39571637828497	-0.236502082498801	easy	-0.159745962996397	-0.79857212609499	1.27364419517435
expensive	0.32476616714178	-2.72876632434145	2.63388382900006	grotesque	-0.159745962996397	-0.79857212609499	1.27364419517435
dangerous	-0.399364907490992	-1.99643031523747	3.18411048793587	pretty	-0.159745962996397	-0.79857212609499	1.27364419517435
happy	-1.03834875947658	2.87239799884985	-1.48126357998888	sweet	-0.159745962996397	-0.79857212609499	1.27364419517435
afraid	-0.958475777978381	2.65144430663063	-1.36732022768204	small	-0.908217045128961	-0.322571845236829	0.974052891853377
big	-2.87659948332055	1.43615386616175	0.335477858532321	cool	-0.399364907490992	1.10476846109609	-0.569716761534184
drunk	-0.878602796480182	2.43049061441141	-1.25337687537520	dumb	-0.399364907490992	1.10476846109609	-0.569716761534184
beautiful	-0.319491925992793	-1.59714425218998	2.54728839034870	lazy	-0.399364907490992	1.10476846109609	-0.569716761534184
precious	-0.319491925992793	-1.59714425218998	2.54728839034870	polite	-0.399364907490992	1.10476846109609	-0.569716761534184
valuable	-0.319491925992793	-1.59714425218998	2.54728839034870	preoccupied	-0.399364907490992	1.10476846109609	-0.569716761534184

# Quantitative analysis: Cluster analysis and adjective classes



- Cluster analysis of top 50 attracted items
- Some clusters corresponding to classes in the (enhanced) version of Dixon's (2004) typology



# Concluding remarks

- Categories based on:
  - Force-dynamic relations between the ADJ- and V-positions
    - Enablement (based on ENABLEMENT image schema)
    - Prevention (based on BLOCKAGE image schema)
  - Zero-anaphoric relations between infinitive clause and co-text
    - Primary participant
    - Secondary participant
    - Condition/factor
- Scalarity of ADJ-position seems to be reflected in construction-lexeme attraction patterns (collexeme analysis)
- Force-dynamic categories seem to be reflected in ADJ-V semantic coherence (covarying collexeme analysis)
- Enablement category seems particularly associated with human propensity adjectives
- Condition/factor category seems particularly associated with the following adjective classes
  - Temperature
  - Atmosphere
  - Time



# Bibliography and references

- Bergen, B. K., & Binsted, K. (2004). The cognitive linguistics of scalar humor. In M. Achard, & S. Kemmer (Eds.), *Language, culture, and mind* (pp. 79-91). Chicago: Chicago University Press.
- Biber, D., Conrad, S., & Reppen, R. (1998). *Corpus linguistics: Investigating language structure and use*. Cambridge: Cambridge University Press.
- Croft, W. A. (2003). Lexical rules vs. constructions: A false dichotomy. In H. Cuyckens, T. Berg, R. Dirven & K. Panther (Eds.), *Motivation in language: Studies in honour of Günter Radden* (pp. 49-68). Amsterdam: John Benjamins.
- Croft, W. A. (2005). Logical and typological arguments for radical construction grammar. In J. Östman (Ed.), *Construction grammars: Cognitive grounding and theoretical extensions* (pp. 273-314). Amsterdam: John Benjamins.
- Davies, M. (2008-2013). *The corpus of contemporary American English (COCA)*. <http://corpus.byu.edu/coca/>
- de Swart, H. (2003). Coercion in a cross-linguistic theory of aspect. In E. J. Francis, & L. A. Michaelis (Eds.), *Mismatch: Form-function incongruity and the architecture of grammar* (pp. 231-258). Stanford: Stanford University Press.
- Dixon, R. M. W. (2004). Adjective classes in typological perspective. In R. M. W. Dixon, & A. Y. Aikhenvald (Eds.), *Adjective classes: A cross-linguistic typology* (pp. 1-47). Oxford: Oxford University Press.
- Fillmore, C. J. (1988). The mechanics of "Construction grammar". *Berkeley Linguistics Society*, 15, 35-55.
- Gries, S. Th. (2009). Coll.analysis 3.2: A program for R for Windows 2.x
- Gries, S. Th., & Stefanowitsch, A. (2004a). Extending collocation analysis: A corpus-based survey on "alternations". *International Journal of Corpus Linguistics*, 9(1), 97-129.
- Gries, S. Th., & Stefanowitsch, A. (2004b). Coracying collexemes in the into-causative. In M. Achard & S. Kemmer (Eds.). *Language, Culture, and Mind*. Chicago: CSLI.
- Hopper, P. (1987). Emergent grammar. *Berkeley Linguistics Society*, 13, 139-157.
- Jensen, K. E. (2013). Custerizer: A program for R for Windows 2.x
- Johnson, M. (1987). *The body in the mind: The bodily basis of meaning, imagination, and reason*. Oxford: University of Chicago Press.
- Kemmer, S., & Barlow, M. (2000). Introduction: A usage-based conception of language. In M. Barlow, & S. Kemmer (Eds.), *Usage-based models of language* (pp. vii-xxviii). Stanford, CA: Stanford University Press.
- Kirk, J. M. (1996). Review of English corpus linguistics: Studies in honour of Jan Svartvik edited by Karin Aijmer and Bengt Altenberg. London: Longman, 1991. *Journal of English Linguistics*, 24(3), 250-258.
- Paradis, C. (2000). It's well weird: Degree modifiers of adjectives revisited: The nineties. In J. M. Kirk (Ed.), *Corpora galore: Analyses and techniques in describing English* (pp. 147-160). Amsterdam: Rodopi.
- Stefanowitsch, A., & Gries, S. T. (2003). Collocations: Investigating the interaction between words and constructions. *International Journal of Corpus Linguistics*, 8(2), 2-43.
- Stefanowitsch, A., & Gries, S. T. (2005). Covarying collexemes. *Corpus Linguistics and Linguistic Theory*, 1(1), 1-43.
- Stefanowitsch, A., & Gries, S. T. (2010). Cluster analysis and the identification of collexeme classes. In S. Rice & J. Newman (Eds.), *Empirical and experimental methods in cognitive/functional research*. Chicago: CSLI.
- Talmy, L. (2000). *Toward a cognitive semantics. vol. 1: Concept structuring systems*. Cambridge, MA: MIT Press.
- Tomasello, M. (2003). *Constructing a language: A usage-based theory of language acquisition*. Cambridge, Mass: Harvard University Press.
- Tummers, J., Heylen, K., & Geeraerts, D. (2005). Usage-based approaches in cognitive linguistics: A technical state of the art. *Corpus Linguistics and Linguistic Theory*, 1(2), 225-261.