

A corpus-based contrastive analysis of the Dutch adjectival -s ending.
Deflexion or refunctionalization?

Dirk Pijpops & Freek Van de Velde

Research Foundation Flanders (FWO)
University of Leuven

WHO

WHAT

WHY

HOW

MOROCCAN-NETHERLANDIC
CHATTERS

Moroccorp corpus

L2/2L1-speakers

Ethnolect

AUTOCHTHONOUS NETHERLANDIC
CHATTERS

ConDiv corpus

L1-speakers

Informal Netherlandic Dutch

WHO

WHAT

WHY

HOW

PARTITIVE GENITIVE

iets leuks

'something fun'

wat bijzonders

'something special'

niets lelijks

'nothing ugly'

iets heel erg interessants

'something very, very interesting'

...

[Indefinite pronoun + Adjectival phrase (-s)]_{Noun phrase}

iets wit

'something white'

niets verkeerd

'nothing wrong'

...

PARTITIVE GENITIVE: CONTRASTIVELY WEIRD

- English: only variant with $-\emptyset$: *something fun, nothing special,...*
- German: $-(e)s$ is part of 'normal' adjectival inflection

etwas Gutes - *iets goeds* - *something good*

zu etwas Gutem - *tot iets goeds* - *to something good*

PARTITIVE GENITIVE

More -Ø

- With color adjectives and assessment adjectives: *verkeerd* 'wrong', *goed* 'good', *beter* 'better', *fout* 'incorrect'
- In Belgium
- In informal language use
- With the pronouns *iets* 'something' and *niets* 'nothing', in Belgium
- In infrequent partitive genitive phrases

WHO

WHAT

WHY

HOW

4 OPTIONS

GENERALIZE THE -S

GENERALIZE THE - \emptyset

USE BOTH, SAME FACTORS

USE BOTH, DIFFERENT FACTORS

4 OPTIONS

GENERALIZE THE -S

GENERALIZE THE - \emptyset

USE BOTH, SAME FACTORS

USE BOTH, DIFFERENT FACTORS

GENERALIZE THE -S

- Ethnolect speakers generalize the most frequent variant
- Ethnolect speakers clean up adjectival inflection
 - -e for attributive prenominal adjectives: e.g. *Het spannende verhaal* 'the suspenseful story'
 - -∅ for attributive prenominal adjectives with indefinite, neuter, singular nouns:
e.g. *Een spannend verhaal* 'a suspenseful story'
 - -∅ for predicative adjectives: e.g. *Het verhaal is spannend* 'the story is suspenseful'
 - -s for attributive postnominal adjectives: e.g. *iets spannends* 'something suspenseful'
 - -∅ frequent for attributive postnominal color & assessment adjectives:
e.g. *iets geel* 'something yellow', *iets verkeerd* 'something wrong'

4 OPTIONS

GENERALIZE THE -S

GENERALIZE THE -Ø

USE BOTH, SAME FACTORS

USE BOTH, DIFFERENT FACTORS

GENERALIZE THE -Ø

- Breakdown of the case system: L2-speakers (Lupyan & Dale 2010, Bentz & Winter 2013)

4 OPTIONS

GENERALIZE THE -S

GENERALIZE THE - \emptyset

USE BOTH, SAME FACTORS

USE BOTH, DIFFERENT FACTORS

USE BOTH, SAME FACTORS

- More -Ø with color adjectives and assessment adjectives: *verkeerd* 'wrong', *goed* 'good', *beter* 'better', *fout* 'incorrect'
 - Grammar rule?
 - Constructional contamination: direct result of the way we process language

USE BOTH, SAME FACTORS

- Constructional contamination

Adverb:

*Of heb ik hier iets **verkeerd** verstaan?*

Or have I here something wrongly understood?

Partitive genitive:

*Als ik **iets verkeerd** gegeten heb,...*

if I something wrong eaten have,...

USE BOTH, SAME FACTORS

- More -Ø with color adjectives and assessment adjectives: *verkeerd* 'wrong', *goed* 'good', *beter* 'better', *fout* 'incorrect'
 - Grammar rule?
 - Constructional contamination: direct result of the way we process language

4 OPTIONS

GENERALIZE THE -S

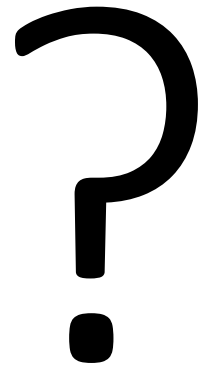
GENERALIZE THE - \emptyset

USE BOTH, SAME FACTORS

USE BOTH, DIFFERENT FACTORS

USE BOTH, OTHER FACTORS

- Exaptation (Lass 1990, Van de Velde & Norde 2016)



4 OPTIONS

GENERALIZE THE $-S$: *CLEAN UP!*

GENERALIZE THE $-\emptyset$: *DEFLECT!*

USE BOTH, SAME FACTORS: *CONTAMINATE!*

USE BOTH, DIFFERENT FACTORS: *REFUNCTIONALIZE!*

WHO

WHAT

WHY

HOW

EXTRACTION

- Pronouns:

iets 'something', *niets* 'nothing', *wat* 'something', *veel* 'a lot', *zoveel* 'so much'

- Adjectives:

aardig 'nice', *apart* 'apart', *belangrijk* 'important', *beter* 'better', *bijzonder* 'particular', *blauw* 'blue', *concreet* 'concrete', *deftig* 'decent', *dergelijk* 'similar', *erg* 'awful', *geel* 'yellow', *gek* 'crazy', *goed* 'good', *groen* 'green', *interessant* 'interesting', *klein* 'small', *lekker* 'tasty', *leuk* 'fun', *mooi* 'beautiful', *nieuw* 'new', *nuttig* 'useful', *oranje* 'orange', *positief* 'positive', *purper* 'purple', *raar* 'weird', *rood* 'red', *spannend* 'exciting', *speciaal* 'special', *verkeerd* 'wrong', *verschrikkelijk* 'horrible', *vreemd* 'weird', *warm* 'warm', *wit* 'white', *zinnig* 'sensible', *zwart* 'black'

(Pijpops & Van de Velde 2014: 9-12)

VARIABLES

- Response variable: presence of -s
- Explanatory variables:
 - Type of Adjective: *assessment adjectives, color adjectives, regular adjectives*
 - Frequency of the phrase, log-transformed
 - Pronoun: *iets* 'something', *niets* 'nothing', *veel* 'a lot', *zoveel* 'so much'
 - Number of syllables of the adjective
 - Number of words of the adjectival phrase
 - Corpus: *Moroccorp, ConDiv*

	Moroccorp	ConDiv
-∅ ending	149	86
-s ending	1464	679

4 OPTIONS

GENERALIZE THE $-S$: *CLEAN UP!*

GENERALIZE THE $-\emptyset$: *DEFLECT!*

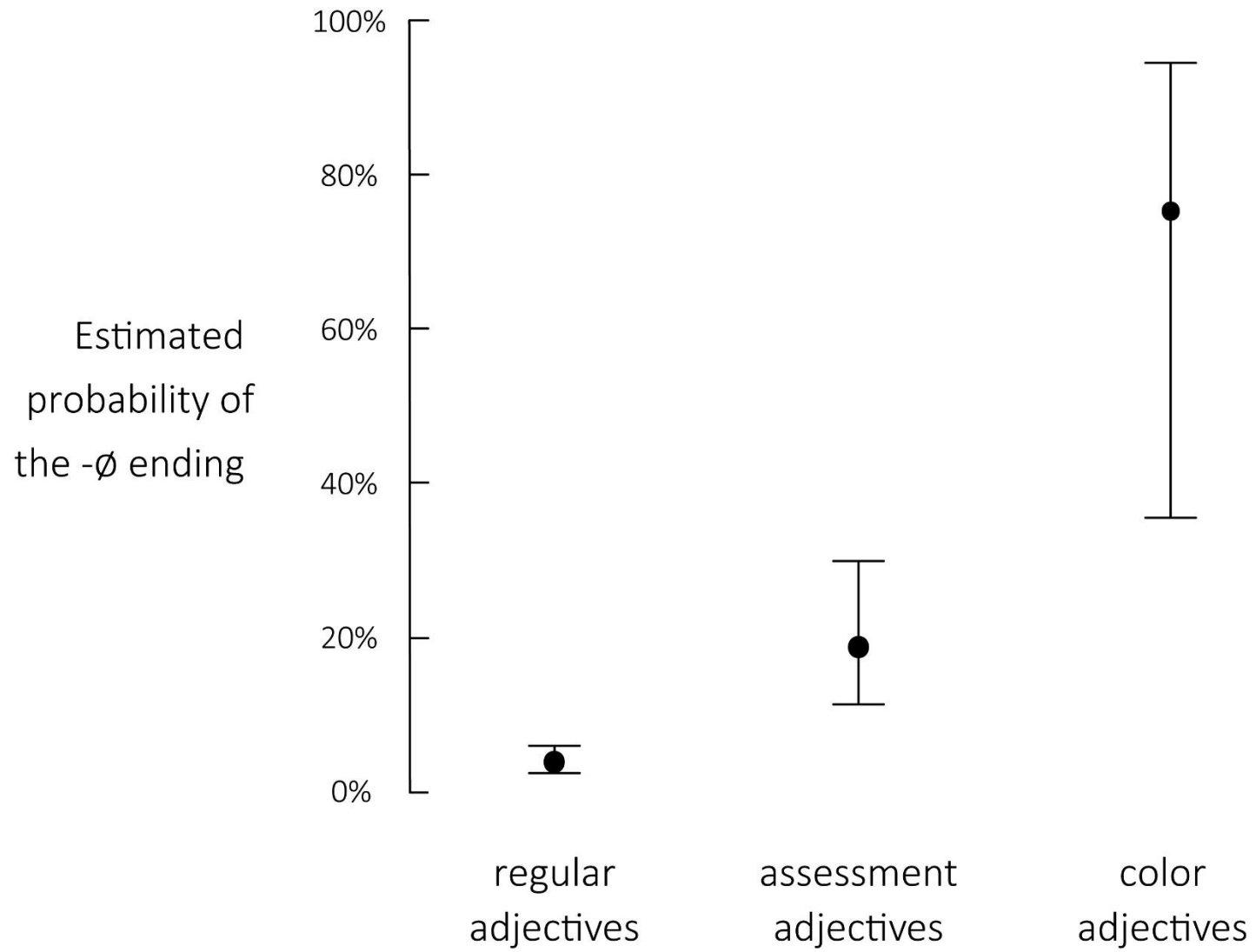
USE BOTH, SAME FACTORS: *CONTAMINATE!*

USE BOTH, DIFFERENT FACTORS: *REFUNCTIONALIZE!*

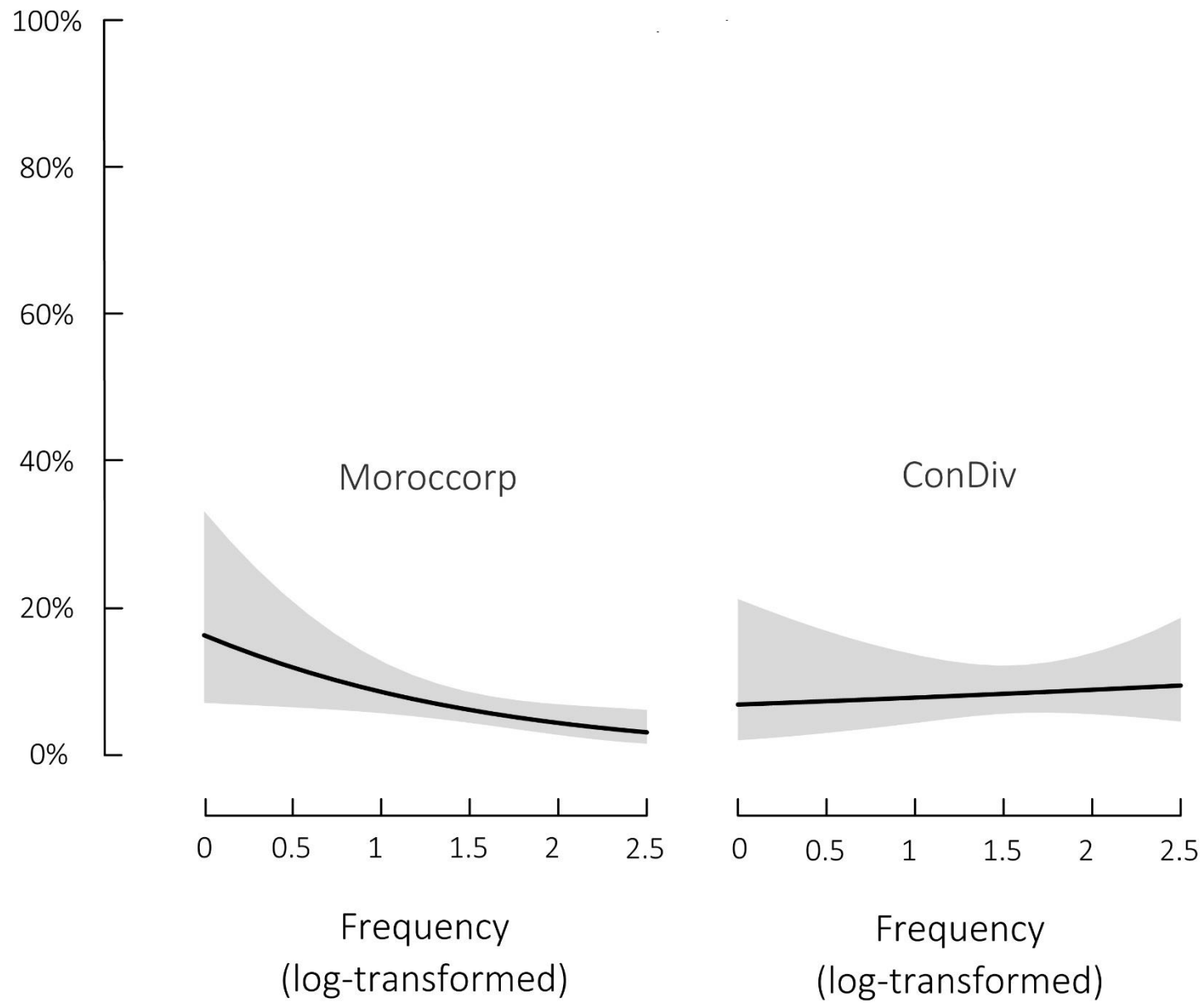
REGRESSION MODEL

- Stepwise variable selection procedure: all explanatory variables & all possible two-way interactions
 - Success level: - \emptyset variant
 - Categorical variables *Type of adjective*, *Pronoun* and *Corpus* implemented with dummy coding
- Added random effect *Phrase*, to control for random lexical preferences
- Model Diagnostics
 - Number of parameters < least frequent response level divided by 20
 - Residual deviance not much larger than degrees of freedom
 - Hosmer-Lemeshow-Cessie goodness-of-fit test: not significant
 - Variance Inflation factors below 4

Predictors	Levels	Estimates	Confidence intervals		P-values
			2,5%	97,5%	
	intercept	-2.09	-2.98	-1.19	< 0.0001
Type of Adjective	<i>regular</i>	Reference level			
	<i>assessment</i>	1.78	1.07	2.48	< 0.0001
	<i>colour</i>	4.34	2.71	5.97	< 0.0001
Frequency		-0.73	-1.32	-0.14	0.0153
Corpus	<i>Moroccorp</i>	Reference level			
	<i>ConDiv</i>	-0.98	-2.28	0.33	0.1416
Interaction Frequency - Corpus	<i>Moroccorp</i>	Reference level			
	<i>ConDiv</i>	0.87	0.15	1.59	0.0175



Estimated probability of the -∅ ending



Predictors	Levels	Estimates	Confidence intervals		P-values
			2,5%	97,5%	
	intercept	-2.09	-2.98	-1.19	< 0.0001
Type of Adjective	<i>regular</i>	Reference level			
	<i>assessment</i>	1.78	1.07	2.48	< 0.0001
	<i>colour</i>	4.34	2.71	5.97	< 0.0001
Frequency		-0.73	-1.32	-0.14	0.0153
Corpus	<i>Moroccorp</i>	Reference level			
	<i>ConDiv</i>	-0.98	-2.28	0.33	0.1416
Interaction Frequency - Corpus	<i>Moroccorp</i>	Reference level			
	<i>ConDiv</i>	0.87	0.15	1.59	0.0175

4 OPTIONS

GENERALIZE THE -S: *CLEAN UP!*

GENERALIZE THE -Ø: *DEFLECT!*

USE BOTH, SAME FACTORS: *CONTAMINATE!*

USE BOTH, DIFFERENT FACTORS: *REFUNCTIONALIZE!*

INTERESTED?

- Contrastive Interlanguage Analysis:

Granger, Sylviane. 2015. Contrastive interlanguage analysis. A reappraisal. *International Journal of Learner Corpus Research* 1(1). 7–24.

- MuPDAR technique:

Gries, Stefan Thomas and Sandra Deshors. 2014. Using regressions to explore deviations between corpus data and a standard/target: two suggestions. *Corpora* 9(1). 109–136.

- 'Chunking' method of language processing:

Dąbrowska, Ewa. 2014. Recycling utterances: A speaker's guide to sentence processing. *Cognitive Linguistics* 25(4). 617–653.

INTERESTED?

- This study:
Pijpops, Dirk and Freek Van de Velde. 2015. Ethnolect speakers and Dutch partitive adjectival inflection. A corpus analysis. *Taal en Tongval* 67(2). 343–371.
- Constructional contamination:
Pijpops, Dirk and Freek Van De Velde. 2016. Constructional contamination: How does it work and how do we measure it? *Folia Linguistica* 50(2).
Pijpops, Dirk, Isabeau De Smet and Freek Van de Velde. Forthcoming. Constructional contamination in morphology and syntax. Four case studies. *Constructions and Frames*.
- Come talk to me

REFERENCES

- Anthony, Laurence. 2011. *AntConc* (Computer Software, version 3.3.3). Tokyo: Waseda University.
- Bates, Douglas, Martin Maechler, Ben Bolker and Steven Walker. 2013. *lme4: Linear mixed-effects models using Eigen and S4*. R package version 1.4.
- Bentz, Christian and Bodo Winter. 2013. Languages with More Second Language Learners Tend to Lose Nominal Case. *Language Dynamics and Change* 3(1). 1–27.
- Bybee, Joan. 2006. From Usage to Grammar: The Mind's Response to Repetition. *Language* 82(4). 711–733.
- Cornips, Leonie and Vincent de Rooij. 2003. Kijk, Levi's is een goeie merk: maar toch hadden ze 'm gedist van je schoenen doen 'm niet. In Jan Stroop (ed.), *Waar gaat het Nederlands naartoe? Panorama van een taal*, 131–142. Amsterdam: Bert Bakker.
- Dąbrowska, Ewa. 2014. Recycling utterances: A speaker's guide to sentence processing. *Cognitive Linguistics* 25(4). 617–653.
- Fox, John, Sanford Weisberg, Michael Friendly, Jangman Hong, Robert Andersen, David Firth and Steve Taylor. 2016. *Effect Displays for Linear, Generalized Linear, and Other Models*. R package version 3.2.
- Granger, Sylviane. 2015. Contrastive interlanguage analysis. A reappraisal. *International Journal of Learner Corpus Research* 1(1). 7–24.
- Grondelaers, Stefan, Katrien Deygers, Hilde Van Aken, Vicky Van den Heede and Dirk Speelman. 2000. Het CONDIV-corpus geschreven Nederlands [The CONDIV-corpus of written Dutch]. *Nederlandse Taalkunde* 5(4). 356–363.
- Lass, Roger. 1990. How to do things with junk: exaptation in language evolution. *Journal of linguistics* 26(1). 79–102.
- Lupyan, Gary and Rick Dale. 2010. Language structure is partly determined by social structure. *PloS one* 5(1). e8559.
- Pijpops, Dirk and Freek Van de Velde. 2014. A multivariate analysis of the partitive genitive in Dutch. Bringing quantitative data into a theoretical discussion. *Corpus Linguistics and Linguistic Theory*. Published online, ahead of print.
- Pijpops, Dirk and Freek Van de Velde. 2015. Ethnolect speakers and Dutch partitive adjectival inflection. A corpus analysis. *Taal en Tongval* 67(2). 343–371.
- Pijpops, Dirk and Freek Van de Velde. 2016. Constructional contamination: How does it work and how do we measure it? *Folia Linguistica* 50(2). 543–581.
- Pijpops, Dirk, Isabeau De Smet and Freek Van de Velde. Constructional contamination in morphology and syntax. Four case studies. *Constructions and Frames*.
- R Core Team. 2014. *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. Vienna.
- Speelman, Dirk, Kris Heylen and Dirk Geeraerts. 2018. *Mixed-Effects Regression Models in Linguistics*. Cham: Springer.
- Speelman, Dirk. 2014. Logistic regression: A confirmatory technique for comparisons in corpus linguistics. In Dylan Glynn & Justyna A. Robinson (eds.), *Corpus Methods for Semantics: Quantitative studies in polysemy and synonymy*, 487–533. (Human Cognitive Processing [HCP]). Amsterdam: John Benjamins.
- Van de Velde, Freek and Fred Weerman. 2014. The resilient nature of adjectival inflection in Dutch. In Petra Sleeman, Freek Van de Velde & Harry Perridon (eds.), *Adjectives in Germanic and Romance*, 113–145. (Linguistik Aktuell/Linguistics Today). Amsterdam: John Benjamins.
- Van de Velde, Freek and Muriel Norde. 2016. Exaptation: taking stock of a controversial notion in linguistics. In Muriel Norde & Freek Van de Velde (eds.), *Exaptation in language change*, 1–35. Amsterdam: John Benjamins.