

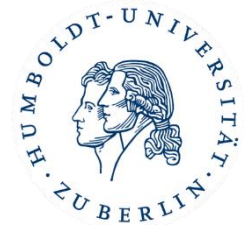
Extravagant ‘fake’ morphemes in Dutch

Morphological productivity, semantic profiles and
categorical flexibility

 UCLouvain

fnrs
LA LIBERTÉ DE CHERCHER

KRISTEL VAN GOETHEM
MURIEL NORDE



1. Introduction

FNRS-CDR J.0211.20 « ACCROSS » (2020-2022)

- ❑ “**Approximative Constructions aCROSS languages: forms, functions and linguistic change**”
- ❑ Collaboration with Muriel Norde (Humboldt University Berlin) and Francesca Masini (University of Bologna)
- ❑ The project aims at establishing a **cross-linguistic inventory of “approximative constructions”**.
- ❑ 3 case studies:
 - Cross-linguistic study of *fake* morphemes in compounding
 - Cross-linguistic study of the prefix *pseudo-*
 - *Soort van* ‘sort of’ as a new approximator in Dutch, compared to English and French


1.1. ‘Fake’ morphemes

- Morphemes that allow ‘privative’ uses: ‘a fake X is not an X’
 - e.g. a *neppistool* ‘fake gun’ is not a gun, although it shares certain properties with a true gun (van der Wouden 2018)
 - “However, we contest that *fake* is an across-the-board privative adjective, since a *fake article*, for instance, is most definitely still an article.” (Cappelle et al. 2018: 9)



1.1. 'Fake' morphemes in Dutch

□ Productive as **compound members**:

- *nepjuwelen* 'fake jewels' < *nep*_{V/N} 'to fool; sham, fake'
- *imitatieleer* 'imitation leather' < *imitatie*_N 'imitation'
- *namaakbont* 'fake fur' < *namaak*_V 'to imitate'
- *kunstgras* 'artificial grass' < *kunst*_N 'art'
- *schijnhuwelijk* 'marriage of convenience' < *schijn*_{V/N} 'to appear; appearance'
- *fopspeen* 'dummy (for baby)' < *fop*_V 'to fool sb.'
- *lokeend* 'decoy duck' < *lok*_V 'to lure sb.'
-  *fake-bericht* 'fake message' < *fake*_{N/V/A}



1.2. Innovative ‘fake’ constructions

- **Productively used as compound members:** *nepdrankjes* ‘fake drinks’, *fopwetenschapper* ‘pseudo-scientist’, ...

- **Orthographic/morphological variation:**
 - *nepherinneringen* vs *nep herinneringen* vs *neppe herinneringen* ‘fake memories’

- **Innovative (“extravagant”) constructions: debonded or adjectival use** (hardly mentioned in the dictionaries)

1.2. Innovative 'fake' constructions

Adjectival use:

(corpus examples taken from nlTenTen14)

- **Predicative use:** *Koop nooit kaviaar op straat; het is vaak **namaak**!* 'Never buy caviar on the streets; it's often fake!'
- **Coordination with an Adj:** *geen **namaak**, maar echte parels* 'not fake, but real pearls'
- **Degree modification:** *Meestal vind ik geuren in huidverzorgingsproducten **te nep** of chemisch ruiken* 'Usually I find perfumes in skin care products smell too fake or chemical'
- **Inflection:** ***neppe** families met hun nog **neppure** plezier* 'fake families and their even faker fun'; *een steeds strakker en **faker** gezicht* 'an ever tighter and faker face'
- **Adverbial use:** *Bent u ook **nep** aardig of echt aardig?* 'Are you also fake nice or really nice?'

1.3. Aims of this study

- ❑ Compare the 8 ‘fake’ morphemes (semantic-distributional profile, morphological profile, productivity)
- ❑ Find out which factors trigger their ‘extravagant’ behavior (i.e. debonding and categorical flexibility)

1.4. Outline

- Debonding as an instance of extravagant morphology

- Case study: ‘fake’ morphemes in Dutch
 - Data and method
 - Semantic-distributional profiles
 - Productivity
 - Debonding and categorical flexibility
 - Relation between productivity, debonding and semantic profile

2. Debonding as extravagant morphology

2.1. Extravagant morphology

- Fortin (2011: ii/175): “**Expressive morphology** of many different languages (...) has the following properties:
 - 1) it is systematically **anomalous** when compared to plain morphology, or the ordinary processes of word-formation and inflection. (...)
 - 2) it is far more **pervasive** than has been traditionally thought.”

- (Partial) overlap between “**evaluative morphology**” (Bauer 1997, Grandi & Körtvélyessy 2015), “**expressive morphology**” (Zwicky & Pullum 1987, Szymanek 1988), “**extragrammatical morphology**” (Dressler & Barbaresi 1994, Dressler & Karpf 1995, Mattiello 2013) and “**extravagant morphology**” (Eitelmann & Haumann 2019)

- Cf. Keller’s (1994) “**maxim of extravagance**”, i.e. “talk in such a way that you are noticed”

2.2. Debonding as extravagant morphology

- ❑ **Debonding** is “a composite change whereby bound morphemes (clitics, affixes, affixoids) in a specific context develop into free morphemes” (Norde 2009: 186).

- ❑ Debonding is **counterdirectional to grammaticalization** (bound > free morpheme)

- ❑ Debonding of intensifying prefixoids is a **productive process of lexical innovation** in Germanic languages, which may lead to the creation of **new intensifying adverbs or evaluative adjectives**
 - For instance: Categorical flexibility of intensifying morphemes (Norde & Van Goethem 2018):
 - Intensifying prefixoid: *Laminaat ligt, maar ik ben ook keikapot ...* ‘Laminate is ready, but I'm exhausted also ...’
 - Intensifying adverb: *Ze komen dus kei te laat (...)* ‘So they are arriving way too late (...)’

2.3. Extravagance and language change

- Debonding challenges Haspelmath (1999)'s claim that grammaticalization is irreversible:

“Grammaticalization is a side effect of the maxim of extravagance, that is, speakers’ use of a unusually explicit formulations in order to attract attention. As these are adopted more widely in the speech community, they become more frequent and are reduced phonologically. I propose that **degrammaticalization is by and large impossible because there is no counteracting maxim of ‘anti-extravagance’ (...).” (Haspelmath 1999: 1043)**

→ However, we consider the changes involved in *degrammaticalization* as more noticeable and extravagant than the pervasive and regular changes of grammaticalization, precisely because *degrammaticalization* runs counter to the “predictable” path of grammaticalization.

3. Case Study: 'fake' morphemes in Dutch

3.1. Data and method

- ❑ Corpus: nITenTen14 (*Sketch Engine*, cf. Killgariff et al. 2014)
- ❑ Random samples of 2000 tokens for each ‘fake’ morpheme, including both bound and free use
- ❑ Manual selection of 500 relevant tokens pro morpheme (388 for *lok*)
- ❑ Annotation of grammatical and semantic properties in Excel
- ❑ Statistical analysis in R (Levshina 2015)

3.2. Semantic-distributional profiles

- Synonymous, interchangeable morphemes?
 - *Die wasabi moet **nep** zijn; hij is veel te goedkoop. Zelfs in Japan zou **namaak**-wasabi veel gebruikt worden.* 'That wasabi must be fake; it's way too cheap. Even in Japan fake wasabi would be widely used.'

- Two types of statistical analyses:
 - Multiple distinctive collexeme analysis
 - Semantic Vector Spaces

Multiple distinctive collexeme analysis

- ❑ Levshina (2015: 248ff.), Stefanowitsch and Gries (2003)
- ❑ compares the observed frequency of a specific slot filler (R1) to the expected frequency of that R1
- ❑ Purpose: compute if specific R1s are attracted to one of the eight 'fake' morphemes

Multiple distinctive collexeme analysis: *nep*

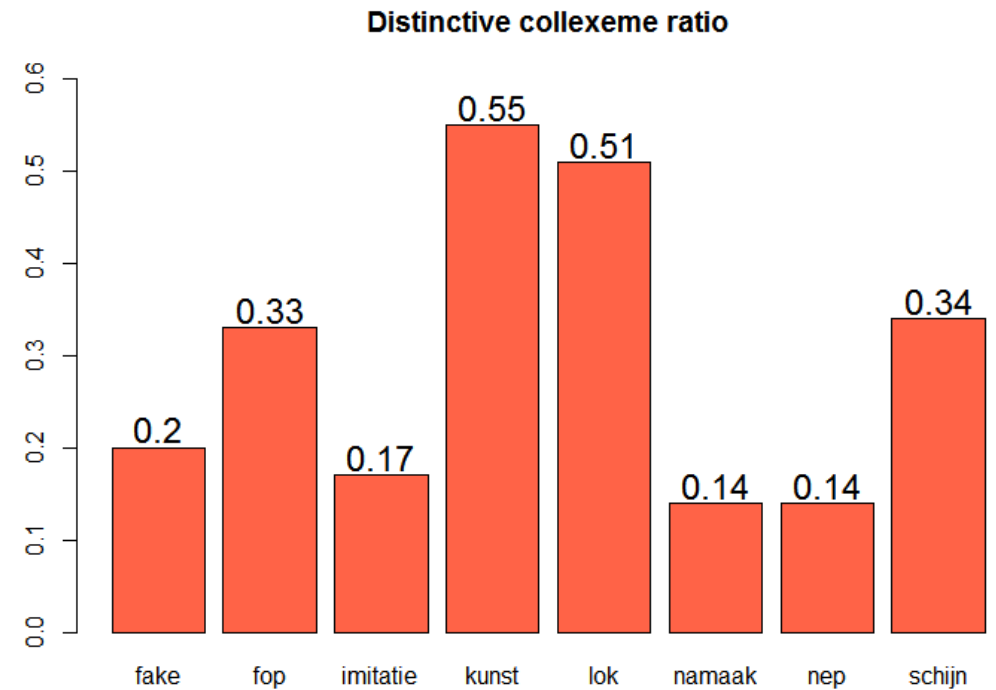
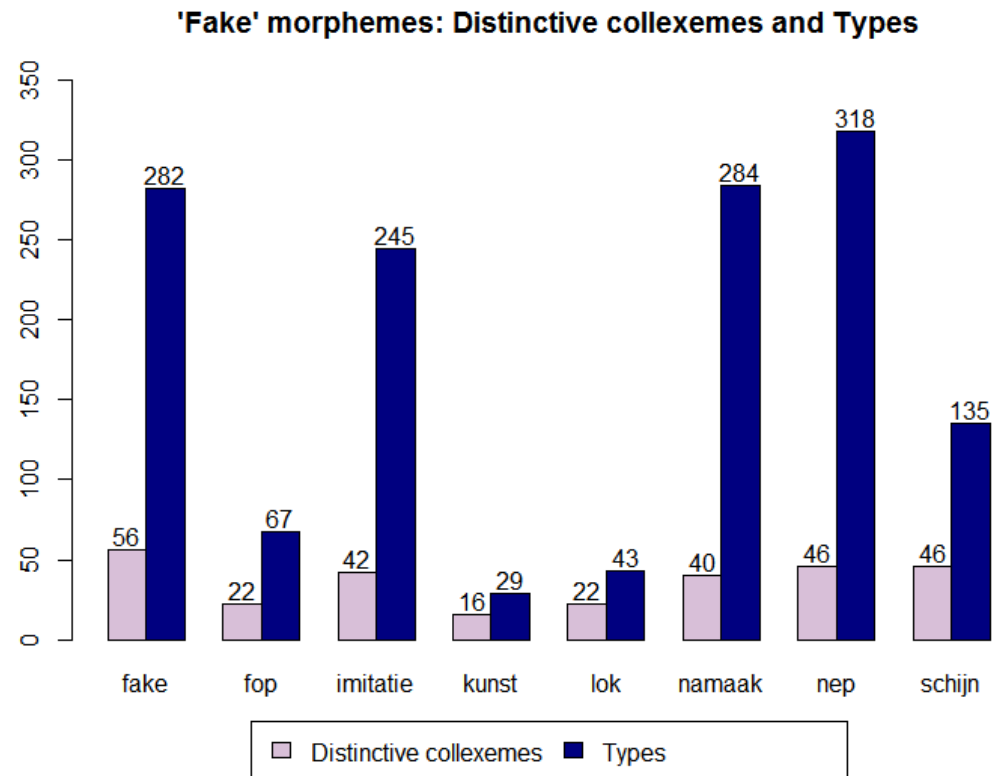
	R1	fake	fop	imitatie	<i>nep</i>	kunst	schijn	lok	namaak	logp
785	wimper	0	0	0	19	1	0	0	0	15.807074
685	nagel	0	0	0	17	8	0	0	0	9.646586
97	geld	1	0	1	9	0	0	0	2	5.398722
600	brief	0	0	0	5	0	0	0	0	4.460817
781	wapen	0	0	0	5	0	0	0	0	4.460817
75	factuur	1	0	0	6	0	0	0	1	4.008330
598	borst	0	0	0	4	0	0	0	0	3.567131
584	beveiligingssoftware	0	0	0	3	0	0	0	0	2.674208
603	brood	0	0	0	3	0	0	0	0	2.674208
648	herinnering	0	0	0	3	0	0	0	0	2.674208
735	snor	0	0	0	3	0	0	0	0	2.674208
759	tiet	0	0	0	3	0	0	0	0	2.674208
774	virusscanner	0	0	0	3	0	0	0	0	2.674208
784	wijn	0	0	0	3	0	0	0	0	2.674208
622	ei	0	0	0	5	0	0	0	4	2.554599
110	haar	2	0	0	4	0	0	0	0	2.484446
163	medicijn	1	0	1	7	0	0	0	8	2.468868
226	site	7	1	0	7	0	0	0	2	2.468868
251	tweet	1	0	0	3	0	0	0	0	2.115966
197	profiel	30	0	0	12	0	0	0	4	1.888459

- Example: distinctive collexemes for *nep*; n = 46 (top 20 shown)
- ($\log p > 1,3 \approx p < 0.05$)
- Various semantic domains: body parts, food, software, ...
- Shared collexemes with *fake* (e.g. *fake/nepprofiel* 'fake profile') and *kunst-* (e.g. *kunst-/nepnagel* 'fake nail')
- Distinctive collexemes (e.g. *nepbrief* 'fake letter')

Multiple distinctive collexeme analysis: top 10 per 'fake' morpheme

	fake	fop	imitatie	nep	kunst	schijn	lok	namaak
1	fur	speen 'teat'	leer 'leather'	wimper 'eyelash'	stof 'material'	heilig 'holy'	vogel 'bird'	product 'product'
2	profile 'profile'	duik 'dive'	bont 'fur'	nagel 'nail'	gras 'grass'	beweging 'movement'	fiets 'bike'	geneesmiddel 'medicine'
3	speech	gesprek 'interview'	suède 'suede'	geld 'money'	mest 'manure'	veiligheid 'security'	eend 'duck'	artikel 'article'
4	account	zwam 'fungus'	leren 'leather'	brief 'letter'	gebit 'teeth'	zekerheid 'certainty'	puber 'adolescent'	sigaret 'cigarette'
5	foto 'photo'	neus 'nose'	spiritualiteit 'spirituality'	wapen 'weapon'	licht 'light'	constructie 'construction'	jood 'Jew'	goederen 'goods'
6	advertentie 'advertisement'	winkel 'shop'	kaviaar 'caviar'	factuur 'invoice'	aas 'bait'	huwelijk 'marriage'	homo 'gay'	versie 'version'
7	tekst 'text'	cadeau 'gift'	parel 'pearl'	borst 'breast'	hars 'resin'	zelfstandigheid 'independence'	auto 'car'	god 'god'
8	naam 'name'	lijn '(telephone) line'	vuurwapen 'firearm'	beveiligingssoftware 'security software'	lens 'lens'	vertoning 'performance'	tiener 'teenager'	Ugg
9	gedoe 'stuff'	mop 'joke'	marmer 'marble'	brood 'bread'	maan 'moon'	zwangerschap 'pregnancy'	duif 'pigeon'	Viagra
10	name	opdracht 'task'	lederen 'leather'	herinnering 'memory'	vezel 'fibre'	oplossing 'solution'	agent 'policeman'	medicijn 'medicine'

Distinctive collexeme ratio



- *Fop, kunst, lok, schijn*: highest distinctive collexeme ratio
→ specific distributional profiles
- *Fake, imitatie, namaak, nep*: lowest distinctive collexeme ratio
→ shared collexemes

Semantic Vector Spaces

- Levshina & Heylen (2014), Levshina (2015: 323ff.)
 - Based on frequency of co-occurrence of target words and context words (for our study: ‘fake’ morpheme + R1)
 - These frequencies are transformed into vectors
 - The cosine of the angle between these vectors can be used as a measure of semantic proximity
- Identifies clusters of (near-)synonyms

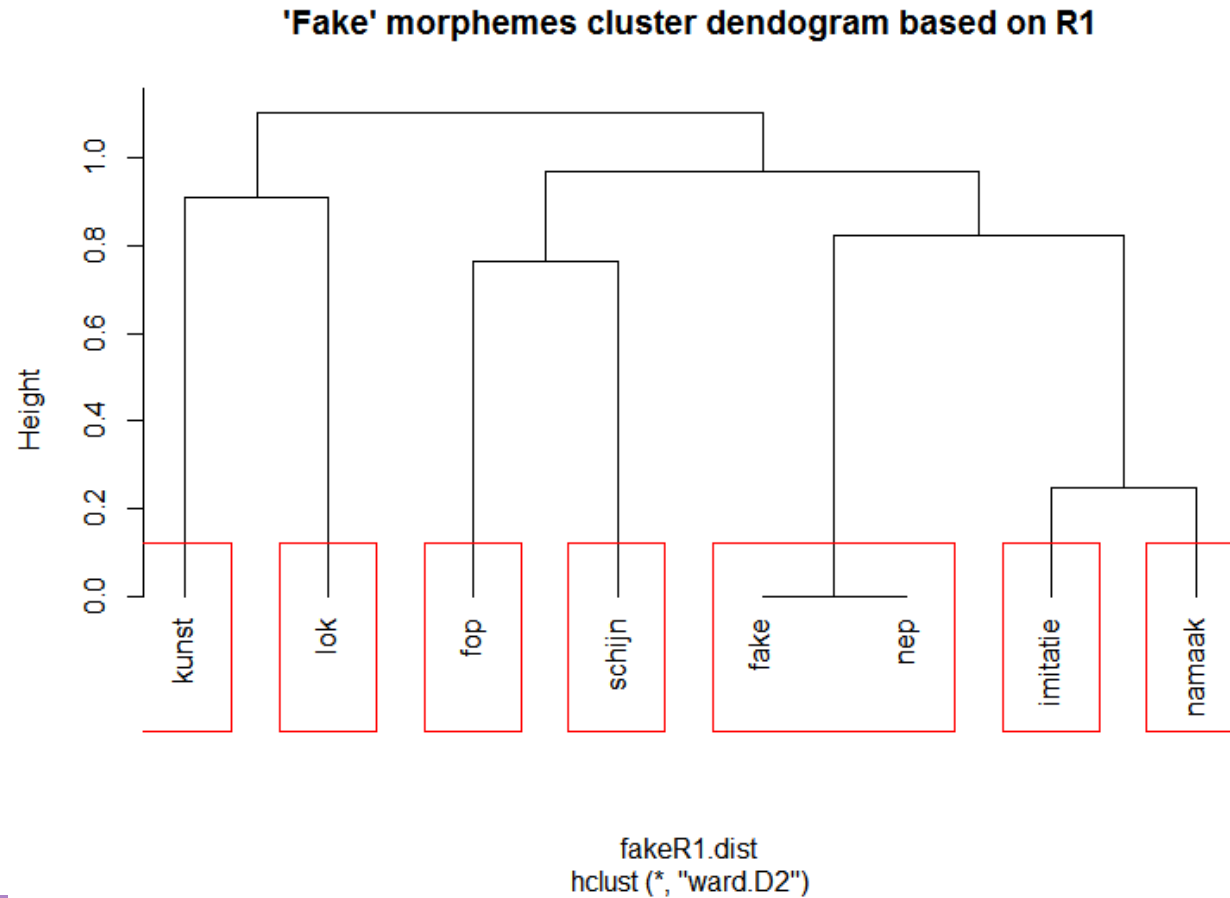
Semantic Vector Spaces: similarity matrix

	fake	fop	imitatie	nep	kunst	schijn	lok	namaak
fake	1.0000	0.0127	0.0060	0.0604	0.0015	0.0212	0.0116	0.0234
fop	0.0127	1.0000	0.0034	0.0292	0.0000	0.0142	0.0000	0.0192
imitatie	0.0060	0.0034	1.0000	0.0325	0.0196	0.0153	0.0000	0.0455
nep	0.0604	0.0292	0.0325	1.0000	0.0142	0.0096	0.0049	0.0525
kunst	0.0015	0.0000	0.0196	0.0142	1.0000	0.0077	0.0055	0.0115
schijn	0.0212	0.0142	0.0153	0.0096	0.0077	1.0000	0.0044	0.0064
lok	0.0116	0.0000	0.0000	0.0049	0.0055	0.0044	1.0000	0.0050
namaak	0.0234	0.0192	0.0455	0.0525	0.0115	0.0064	0.0050	1.0000

Scores range from 0 (no overlap at all between R1s) to 1 (complete overlap between R1s)

- *Fake/nep*: highest similarity scores
- *Nep/namaak, imitatie/namaak*: medium similarity scores
- *fop/kunst, fop/lok, imitatie/lok, kunst/fop*: no overlap at all

Semantic Vector Spaces: cluster dendrogram



Semantic-distributional profiles: summary

- ❑ ***Fake* and *nep***: most similar semantic-distributional profiles, even if they have specific collexemes
 - *Een echte of een **fake kerstboom**? Ik heb samen met mijn vriend een **nep kerstboom** gekocht.* ‘A real or a fake Christmas tree? Together with my friend I bought a fake Christmas tree.’
 - English loanwords more often modified by *fake*: *fake fur, account, speech, ...*
- ❑ ***Imitatie* and *namaak***: also comparable distributional-semantic profiles, but comparison distinctive collexemes shows some subtle differences:
 - *imitatie* primarily refers to faithful copies of expensive natural materials (leather, fur, caviar, marble, etc.), e.g. *imitatiebont* ‘imitation fur’
 - *namaak* more often has a connotation of fraud (medicine, article, cigarettes, or brand names such as Uggs), e.g. *namaaksigaret* ‘counterfeit cigarette’

Semantic-distributional profiles: summary

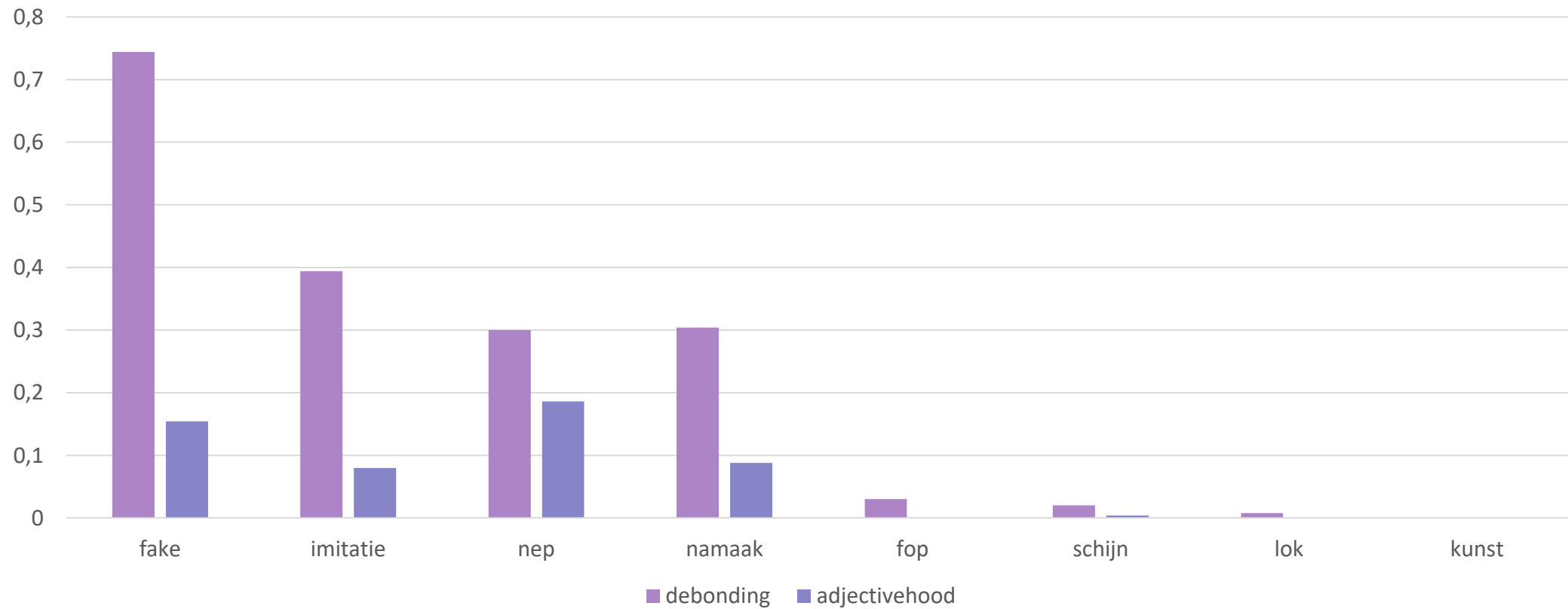
- The other ‘fake’ morphemes have more specific semantic-distributional profiles:
 - **Kunst-X**: “artificial, synthetic, not natural X”
e.g. *kunststof* ‘synthetic material’, *kunstgras* ‘artificial turf’
 - **Schijn-X**: “apparent X, superficial resemblance”, often pejorative connotation
e.g. *schijnheilig* ‘hypocritical’, *schijnveiligheid* ‘false safety’
 - **Fop-X**= “fake X in order to fool sb., intended as a joke”
e.g. *fopspeen* ‘dummy’, *fopgesprek* ‘joke interview’
 - **Lok-X** = “fake X in order to lure sb., intended as a trap”
e.g. *lokvogel* ‘decoy (bird)’, *lokfiets* ‘bait bike’

3.3. Debonding and categorical flexibility

- **Debonding ratio:** purely orthographic criterion
 - e.g. *nep herinnering*

- **Adjectivehood ratio:** morpho-syntactic criterion (predicative use, gradation, inflection, coordination with A, adverbial use)
 - e.g. *een neppe herinnering*

Debonding and adjectivehood ratios



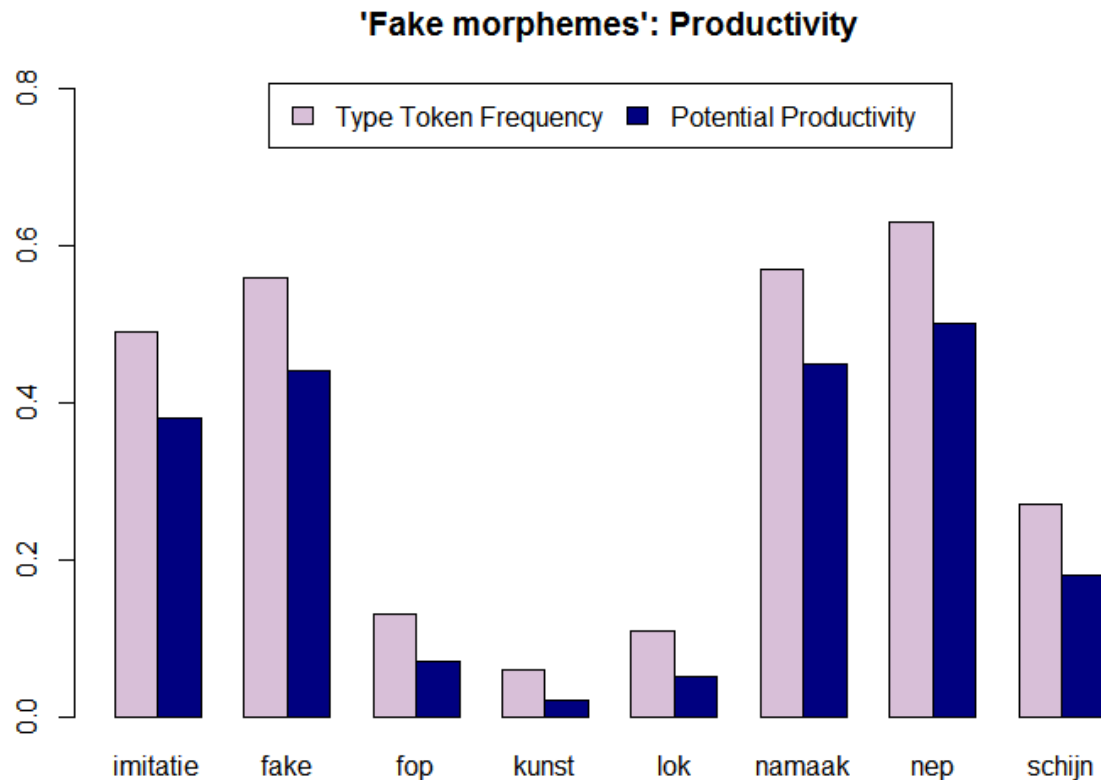
Debonding and categorical flexibility: summary

- ❑ **Fop, kunst, lok, schijn**: mostly bound, no signs of adjectivehood
 - e.g. *schijnbeweging* ‘feint (movement)’, *kunstgebit* ‘false teeth’
- ❑ **Imitatie, namaak**: relatively high degree of debonding, but few other signs of adjectivehood (e.g. no inflection)
 - Longer compounds: e.g. *imitatie sneeuwkettingen* ‘imitation snow chains’
 - Combination with brand names: e.g. *namaak Gucci*
- ❑ **Fake** has the highest degree of debonding (but cf. English spelling), but not of adjectivehood (inflection difficult to determine)
- ❑ **Nep**: high degree of adjectivehood, including inflection and gradation

→ Link between semantic and morphological profile

→ Also a correlation with productivity?

3.4. Productivity



Most productive morphemes: *nep*, *fake*, *namaak*, *imitatie*

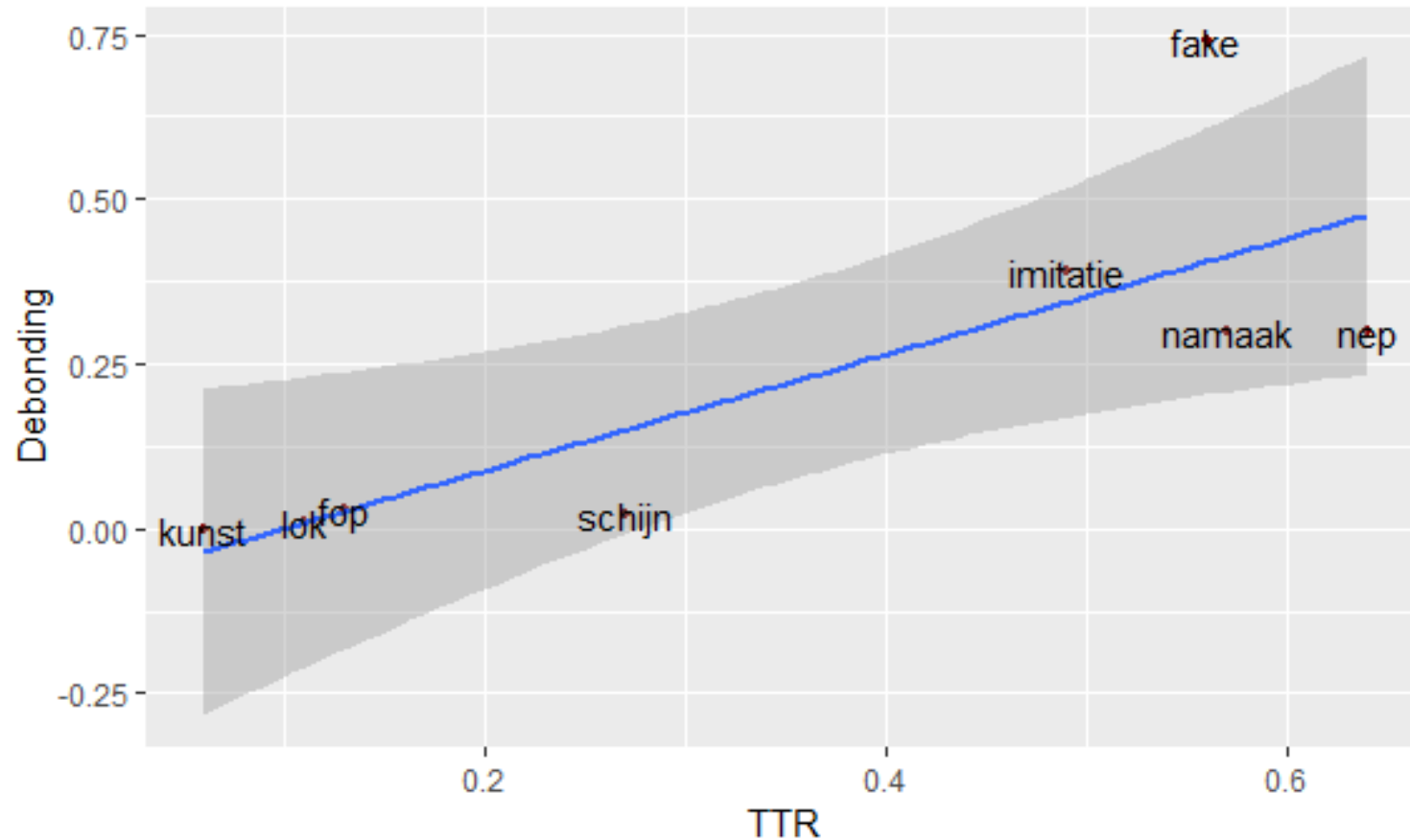
- Hapax legomena: e.g. *nepzwanger* 'falsely pregnant', *namaakwielrenner* 'fake (amateurish) cyclist'

Least productive morphemes: *fop*, *kunst*, *lok*, *schijn*

3.5. Productivity and debonding

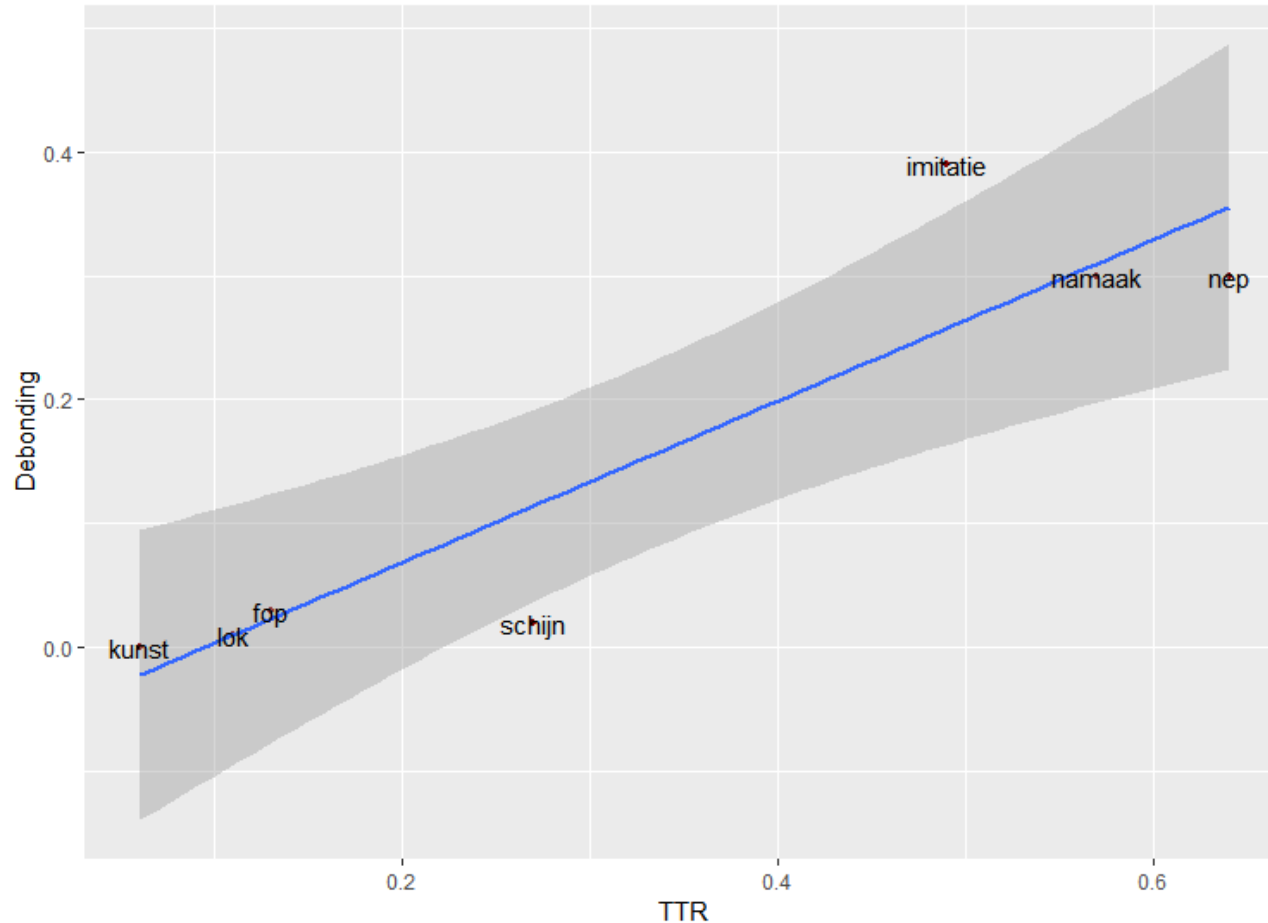
- ❑ Does increasing productivity correlate with increasing debonding ratio?
- ❑ Method: Correlation analysis with linear regression modelling (Levshina 2015: 115ff., 139ff.)

Productivity and debonding: all



- Debonding ratio: response variable
- TTR: explanatory variable
 $p = 0.01815 (*)$
Adjusted R^2 : 0.6334
cor (Pearson's) = 0.796

Productivity and debonding: *fake* excluded



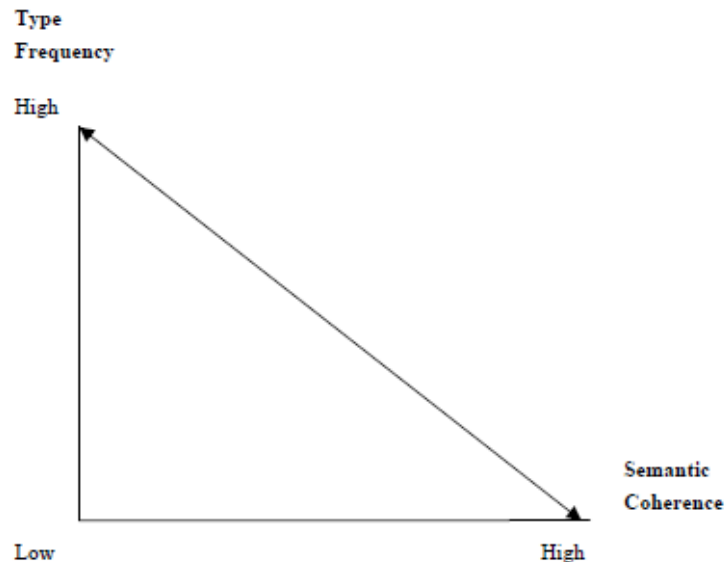
Significant correlation
between TTR and
debonding ratio

- $p = 0.004396 (**)$
- Adjusted R^2 : 0.8288
- $\text{cor (Pearson's)} = 0.910$

4. Discussion and conclusion

Productivity and debonding

- How to explain the correlation between **productivity** and **debonding**?
- Barðdal's (2008): inverse correlation between type frequency and semantic coherence of a construction



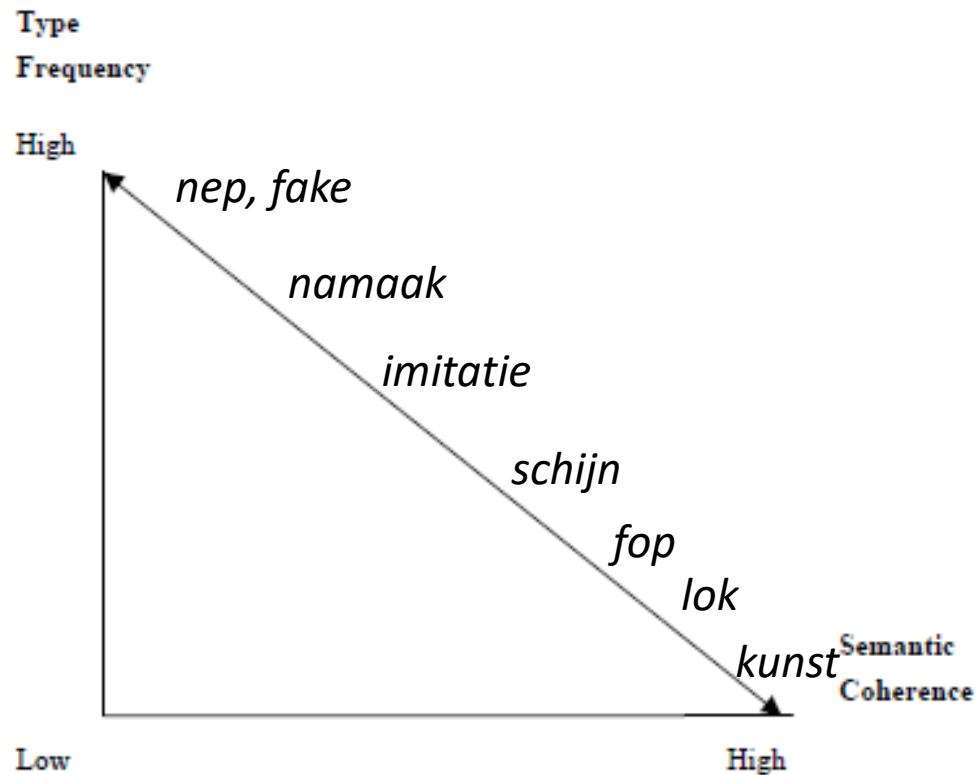
“This means that the higher the type frequency of a construction, the lower the degree of semantic coherence is needed for a construction to be productive. Conversely, the lower the type frequency of a construction, the higher degree of semantic coherence is needed for a construction to be extendable.”
(Barðdal's 2008: 35)

Productivity and debonding

Applied to our case study:

- ***Nep, fake, namaak, imitatie***: ‘fake’ morphemes with relatively low semantic coherence (neutral ‘fake’ semantics, low distinctive collexeme ratio) and high type frequency
 - Low semantic coherence may result in lower morphological cohesion and trigger debonding (and possibly adjectival reanalysis)
- ***Fop, kunst, lok, schijn***: ‘fake’ morphemes with high semantic coherence (more specialized semantics, high distinctive collexeme ratio) and low type frequency
 - High semantic coherence may correlate with stronger morphological cohesion and prevent debonding

A constructionist view



Schematic constructions
(constructions with low semantic coherence and high type frequency)

Semi-schematic constructions

Substantive constructions
(semantically coherent compounds with limited number of types)

SYNTAX

DEBONDING

LEXICAL MORPHOLOGY

Conclusions

- ❑ Debonding of ‘fake’ morphemes is an instance of extravagant morphology (\leftrightarrow Haspelmath 1999)
- ❑ Dutch ‘fake’ morphemes do not show this extravagant behavior to the same extent
- ❑ Extravagant behavior triggered by low degree of semantic coherence and high degree of productivity
- ❑ A cline of ‘fake’ morphemes:
 - Upper part of the cline (*nep, fake, namaak, imitatie*): morphemes with high type frequency and low semantic coherence \rightarrow debonding and possibly categorical flexibility (e.g. *neppe/valse herinneringen* ‘fake/false memories’)
 - Lower part of the cline (*schijn, fop, lok, kunst*): morphemes with low type frequency and high semantic coherence \rightarrow morphologically cohesive compounds (e.g. *fopspeen* ‘dummy’, *kunststof* ‘synthetic’)

References

- Barðdal, Jóhanna (2008), *Productivity: Evidence from case and argument structure in Icelandic*, Amsterdam / Philadelphia: John Benjamins.
- Battefeld, Malte, Torsten Leuschner & Gudrun Rawoens. (2018). ‘Evaluative Morphology’ in German, Dutch and Swedish: Constructional Networks and the Loci of Change. in Kristel Van Goethem, Muriel Norde, Evie Coussé & Gudrun Vanderbauwhede (Eds.) *Category change from a constructional perspective*. 229-262.
- Bauer, Laurie. (1997). Evaluative morphology: In search of universals. *Studies in Language* 21(3). 533–575.
- Cappelle, Bert, Pascal Denis, and Manuela Keller (2018), Facing the facts of fake: a distributional semantics and corpus annotation approach, *Yearbook of the German Cognitive Linguistics Association* 6(1), 9-42.
- Dressler, Wolfgang Ulrich & Lavinia Merlini Barbaresi. (1994). *Morphopragmatics. Diminutives and intensifiers in Italian, German, and other languages*. Berlin & New York: Mouton de Gruyter.
- Dressler, Wolfgang Ulrich & Annemarie Karpf. (1995). The theoretical relevance of Pre- and Protomorphology in language acquisition. In Geert Booij & Jaap van Marle (eds.), *Yearbook of Morphology*, 99-122. Dordrecht: Kluwer.
- Eitelmann, Matthias & Dagmar Haumann. (2019). Extravagant morphology. Workshop organized at the *52nd Annual Meeting of the Societas Linguistica Europaea (SLE 2019)*, 21-24 August 2019, Leipzig University, Germany.
- Fortin, Antonio (2011) *The Morphology and Semantics of Expressive Affixes*. PhD thesis University of Oxford. Online available at <http://ling.auf.net/lingbuzz/001365/v1.pdf>

References

- Grandi, Nicola & Lívia Körtvélyessy (eds). (2015). *Edinburgh Handbook of Evaluative Morphology*. Edinburgh: Edinburgh University Press.
- Haspelmath, Martin (1999). Why is grammaticalization irreversible? *Linguistics* 37(6), 1043–1068.
- Keller, Rudi (1994). *Language Change: The Invisible Hand in Language* (translation of Keller 1990). London: Routledge.
- Kilgarriff, Adam, Vít Baisa, Jan Bušta, Miloš Jakubíček, Vojtěch Kovář, Jan Michelfeit, Pavel Rychlý, and Vít Suchomel (2014), The Sketch Engine: ten years on. *Lexicography* 1, 7-36.
- Körtvélyessy, Lívia. (2014). Evaluative morphology. In Rochelle Lieber and Pavol Štekauer, *The Oxford Handbook of Derivational Morphology*, 296-316. Oxford: Oxford University Press.
- Levshina, Natalia. (2015). How to do Linguistics with R. Data exploration and statistical analysis. Amsterdam/Philadelphia: John Benjamins Publishing Company.
- Mattiello, Elisa. 2013. *Extra-grammatical Morphology in English: Abbreviations, Blends, Reduplicatives, and Related Phenomena*. Berlin & Boston: Mouton de Gruyter
- Maechler, Martin, Rousseeuw, Peter, Struyf, Anja, Hubert, Mia & Kurt Hornik. (2019). cluster: Cluster Analysis Basics and Extensions. R package version 2.0.9.

References

- Norde, Muriel (2009). *Degrammaticalization*. Oxford: Oxford University Press.
- Norde, Muriel & Kristel Van Goethem (2018), Debonding and Clipping of Prefixoids in Germanic: Constructionalization or Constructional Change? in G. Booij (Ed.), *The Construction of Words (Studies in Morphology 4)*, Cham etc.: Springer, 475-518.
- R Core Team (2016). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL <https://www.R-project.org/>
- Stefanowitsch, Anatol & Stefan Th. Gries. (2003). Collostructions: Investigating the interaction of words and constructions. *International Journal of Corpus Linguistics* 8(2). 209-243.
- Van Goethem, Kristel & Matthias Hüning (2015). From noun to evaluative adjective: conversion or debonding? Dutch top and its equivalents in German. *Journal of Germanic linguistics* 27(4). 366-409.
- Wouden, Ton van der (2018). *nep-*. *Taalportaal*. Retrieved from <http://www.taalportaal.org/taalportaal/topic/pid/topic-13998813294731184>. (accessed 26 March 2019).
- Zwicky, Arnold M. & Geoffrey K. Pullum. (1987). Plain morphology and expressive morphology. General session and parasession on grammar and cognition. In Jon Aske, Natasha Beery, Laura Michaelis & Hana Filip (eds.), *Proceedings of the Thirteenth Annual Meeting of the Berkeley Linguistics Society*, 330-340. Berkeley: Berkeley Linguistics Society.

References

- Sketch Engine: <http://www.sketchengine.eu>
- Etymological dictionaries:
 - <http://www.etymonline.com>
 - <http://gtb.inl.nl>
 - <http://www.etymologiebank.nl/>