Nominalizations with the suffixes -*ee* and -*ation* A distributional semantic analysis

Viktoria Schneider & Ingo Plag

InSemantiC22

23/11/2022

Introduction: eventuality-related nominalizations

Introduction: eventuality-related nominalizations

- What are eventuality-related nominalizations?

Introduction: eventuality-related nominalizations

- What are eventuality-related nominalizations?

(1) *employee*, *trainee*
→ participant reading

Introduction: eventuality-related nominalizations

- What are eventuality-related nominalizations?

(1) *employee*, *trainee*
→ participant reading

(2) *Markham sets down the rules about park befoulment*. (Plag et al. 2018: 474)
→ whole eventuality reading

Introduction: eventuality-related nominalizations

- What are eventuality-related nominalizations?

  (1) *employee*, *trainee*
  → participant reading

  (2) *Markham sets down the rules about park befoulment*. (Plag et al. 2018: 474)
  → whole eventuality reading

- Semantic representation provides eventualities and participants for word formation process (e.g., Plag et al. 2018, Kawaletz 2021, Schneider 2023)

Introduction: eventuality-related nominalizations

- What are eventuality-related nominalizations?

  (1) *employee*, *trainee*
  → participant reading

  (2) *Markham sets down the rules about park befoulment*. (Plag et al. 2018: 474)
  → whole eventuality reading

- Semantic representation provides eventualities and participants for word formation process (e.g., Plag et al. 2018, Kawaletz 2021, Schneider 2023)

- Research tends to focus on deverbal nominalizations (e.g., Barker 1998; Alexiadou 2010; Kawaletz & Plag 2015; Plag et al. 2018; Kawaletz 2021)

Introduction: eventuality-related nominalizations

- What are eventuality-related nominalizations?

  (1) *employee*, *trainee*
  → participant reading

  (2) *Markham sets down the rules about park befoulment*. (Plag et al. 2018: 474)
  → whole eventuality reading

- Semantic representation provides eventualities and participants for word formation process (e.g., Plag et al. 2018, Kawaletz 2021, Schneider 2023)

- Research tends to focus on deverbal nominalizations (e.g., Barker 1998; Alexiadou 2010; Kawaletz & Plag 2015; Plag et al. 2018; Kawaletz 2021)

- Many nominalizing suffixes also attach to non-verbal bases (e.g., Plag 1999, 2004; Bauer et al. 2013, Schneider 2023)

Introduction: Distributional Semantics

Introduction: Distributional Semantics

- Distributional Semantics useful for analysis of nominalizations
  (e.g., Lapesa et al. 2018; Wauquier et al. 2018; Huyghe & Wauquier 2020)

Introduction: Distributional Semantics

- Distributional Semantics useful for analysis of nominalizations
  (e.g., Lapesa et al. 2018; Wauquier et al. 2018; Huyghe & Wauquier 2020)

- Difference in meaning = difference in distribution

Introduction: Distributional Semantics

- Distributional Semantics useful for analysis of nominalizations
  (e.g., Lapesa et al. 2018; Wauquier et al. 2018; Huyghe & Wauquier 2020)

- Difference in meaning = difference in distribution

- Word vector: computed by list of words in context of target word

Introduction: Distributional Semantics

- Distributional Semantics useful for analysis of nominalizations
  (e.g., Lapesa et al. 2018; Wauquier et al. 2018; Huyghe & Wauquier 2020)

- Difference in meaning = difference in distribution

- Word vector: computed by list of words in context of target word

- Distance between vectors = semantic similarity
  - High distance → unsimilar
  - Low distance → similar

Introduction: Distributional Semantics

Introduction: Distributional Semantics

Dress: *The dress was in the closet, with the skirt and the t-shirt*

Banana: *The banana tasted as delicious as the apple and the peach*

Introduction: Distributional Semantics

Dress: *The dress was in the closet, with the skirt and the t-shirt*

Banana: *The banana tasted as delicious as the apple and the peach*

	skirt	t-shirt	apple	peach
dress	45	26	3	0
banana	0	2	34	18

Introduction: Distributional Semantics

Dress: *The dress was in the closet, with the skirt and the t-shirt*

Banana: *The banana tasted as delicious as the apple and the peach*

	skirt	t-shirt	apple	peach
dress	45	26	3	0
banana	0	2	34	18

Introduction: Distributional Semantics

Dress: *The dress was in the closet, with the skirt and the t-shirt*

Banana: *The banana tasted as delicious as the apple and the peach*

	skirt	t-shirt	apple	peach
dress	45	26	3	0
banana	0	2	34	18

Introduction: Distributional Semantics

Dress: *The dress was in the closet, with the skirt and the t-shirt*

Banana: *The banana tasted as delicious as the apple and the peach*

	skirt	t-shirt	apple	peach
dress	45	26	3	0
banana	0	2	34	18

Introduction: Distributional Semantics

Dress: *The dress was in the closet, with the skirt and the t-shirt*

Banana: *The banana tasted as delicious as the apple and the peach*

	skirt	t-shirt	apple	peach
dress	45	26	3	0
banana	0	2	34	18

word vectors

dimensions

- Word vector represents a word‘s semantics

Introduction: Distributional Semantics

Dress: *The dress was in the closet, with the skirt and the t-shirt*

Banana: *The banana tasted as delicious as the apple and the peach*

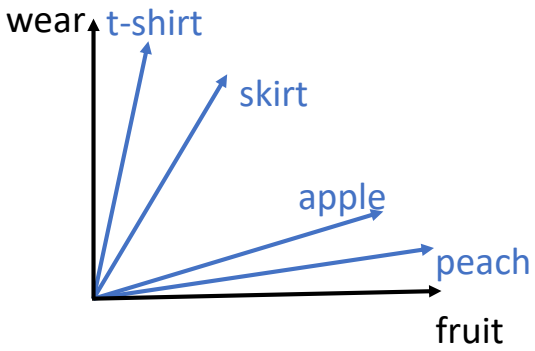
	skirt	t-shirt	apple	peach
dress	45	26	3	0
banana	0	2	34	18

word vectors

dimensions

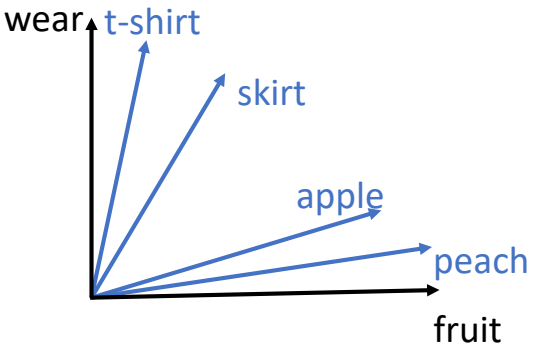
- Word vector represents a word‘s semantics
- Usually more dimensions: for example 300 dimensions

Introduction: Distributional Semantics



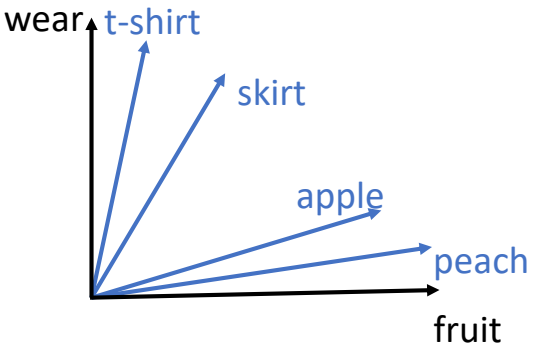
Introduction: Distributional Semantics

Dress: *The dress was in the closet, with the skirt and the t-shirt*



Introduction: Distributional Semantics

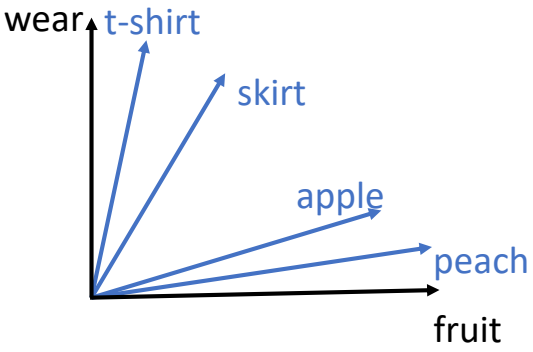
Dress: *The dress was in the closet, with the skirt and the t-shirt*

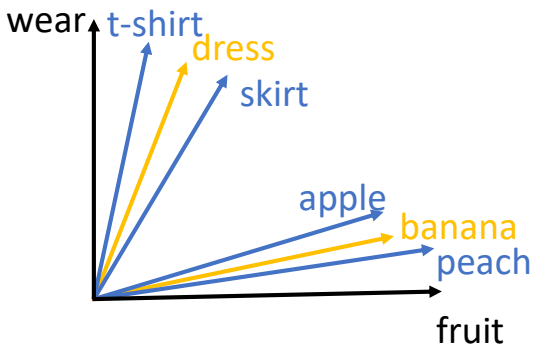


Banana: *The banana tasted as delicious as the apple and the peach*

Introduction: Distributional Semantics

Dress: *The dress was in the closet, with the skirt and the t-shirt*





Banana: *The banana tasted as delicious as the apple and the peach*

Nominalizations & Distributional Semantics

- Deverbal and non-deverbal eventuality-related nominalizations

Nominalizations & Distributional Semantics

- Deverbal and non-deverbal eventuality-related nominalizations

- Are there differences in the semantics of deverbal and denominal nominalizations as the same suffix operates on different word classes?

Nominalizations & Distributional Semantics

- Deverbal and non-deverbal eventuality-related nominalizations

- Are there differences in the semantics of deverbal and denominal nominalizations as the same suffix operates on different word classes?

  → use Distributional Semantics to find out

Nominalizations & Distributional Semantics

- Deverbal and non-deverbal eventuality-related nominalizations

- Are there differences in the semantics of deverbal and denominal nominalizations as the same suffix operates on different word classes?

  → use Distributional Semantics to find out

- Focus on -*ee* and -*ation*

Research questions

Research questions

- How similar are the meanings of a derivative and its base word?

Research questions

- How similar are the meanings of a derivative and its base word?
  - How similar are the meanings of **denominal** derivatives and their **nominal** base words?
  - How similar are the meanings of **deverbal** derivatives and their **verbal** base words?

Research questions

- How similar are the meanings of a derivative and its base word?
  - How similar are the meanings of **denominal** derivatives and their **nominal** base words?
  - How similar are the meanings of **deverbal** derivatives and their **verbal** base words?

- Which factors influence the similarity between base and derivative?

Research questions

- How similar are the meanings of a derivative and its base word?
  - How similar are the meanings of **denominal** derivatives and their **nominal** base words?
  - How similar are the meanings of **deverbal** derivatives and their **verbal** base words?

- Which factors influence the similarity between base and derivative?

- Do we find differences regarding different suffixes, -*ee* and -*ation*?

Hypotheses

Hypotheses

- Base and derivative similar
  - Eventive elements for word formation process already in base
    (e.g., Plag et al. 2018, Kawaletz 2021, Schneider 2023)

Hypotheses

- Base and derivative similar
  - Eventive elements for word formation process already in base
    (e.g., Plag et al. 2018, Kawaletz 2021, Schneider 2023)

- Verbal bases more similar to their derivatives than nominal bases to their derivatives
  - Verbs clearly eventive (e.g., Van Valin & LaPolla 1997; Haspelmath 2001; Szabó 2015)
  - Word formation process more straightforward
  - Eventive elements more easily identifiable for word formation process

Method

- FastText *Common Crawl* subword model (Bojanowski et al. 2016, Mikolov et al. 2018)
  - 2 million pre-trained word vectors
  - Contains subword information → *n*-grams

Method

- FastText *Common Crawl* subword model (Bojanowski et al. 2016, Mikolov et al. 2018)
  - 2 million pre-trained word vectors
  - Contains subword information → *n*-grams

	#dr	dre	res	ess	ss#	#ba	ban	ana	nan	an#
dress	1	1	1	1	1	0	0	0	0	0
banana	0	0	0	0	0	1	1	2	1	1

Method

- FastText *Common Crawl* subword model (Bojanowski et al. 2016, Mikolov et al. 2018)
  - 2 million pre-trained word vectors
  - Contains subword information → *n*-grams

	#dr	dre	res	ess	ss#	#ba	ban	ana	nan	an#
dress	1	1	1	1	1	0	0	0	0	0
banana	0	0	0	0	0	1	1	2	1	1

- create new vectors based on distributional criteria and *n*-grams → also for low frequent ones due to subword information

Method

- FastText *Common Crawl* subword model (Bojanowski et al. 2016, Mikolov et al. 2018)
  - 2 million pre-trained word vectors
  - Contains subword information → *n*-grams

	#dr	dre	res	ess	ss#	#ba	ban	ana	nan	an#
dress	1	1	1	1	1	0	0	0	0	0
banana	0	0	0	0	0	1	1	2	1	1

- create new vectors based on distributional criteria and *n*-grams → also for low frequent ones due to subword information

- Compare cosine similarities of denominal/deverbal derivatives and their nominal/verbal bases

Method

Method

- Beta regression models to determine which factors influence the cosine similarity
  - Dependent variable: cosine similarity between base and derivative, range of [0,1]

Method

- Beta regression models to determine which factors influence the cosine similarity
  - Dependent variable: cosine similarity between base and derivative, range of [0,1]

Variables of interest	Expectation
Relative frequency of base/derivative	Higher relative frequency leads to higher segmentability (e.g., Hay & Baayen 2003) → more transparent → higher cosine similarity
Word class of base	Verbal bases more similar to derivatives due to clearer eventuality
Polysemy of base	Higher polysemy of base leads to decrease of cosine similarity

biographee

debtee

The suffix -*ee*

46 denominal 312 deverbal

mentee

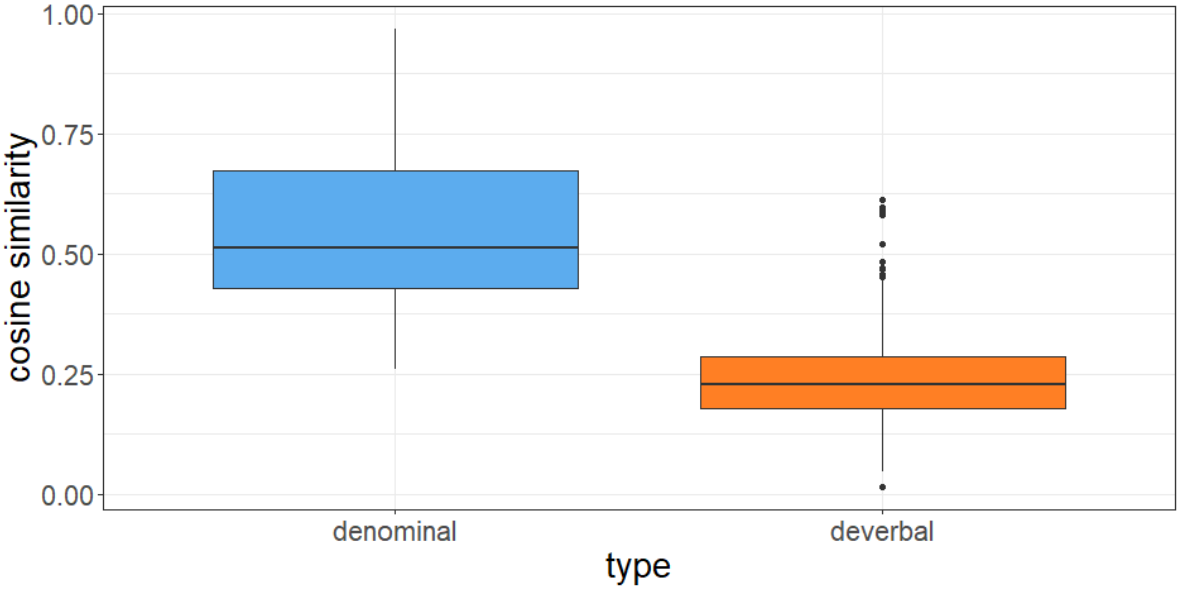
covenantee

tutee

Data from COCA (Davies 2008) and BNC (Davies 2004)

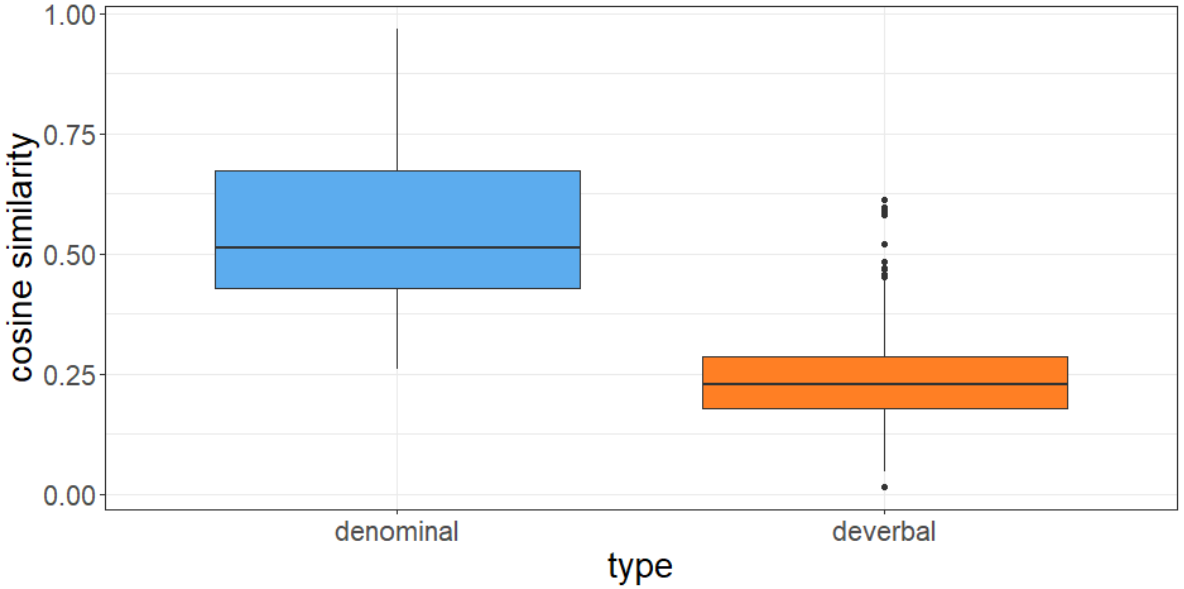
Similarity of nominal and verbal bases and derivatives for -*ee*

Similarity of nominal and verbal bases and derivatives for -*ee*



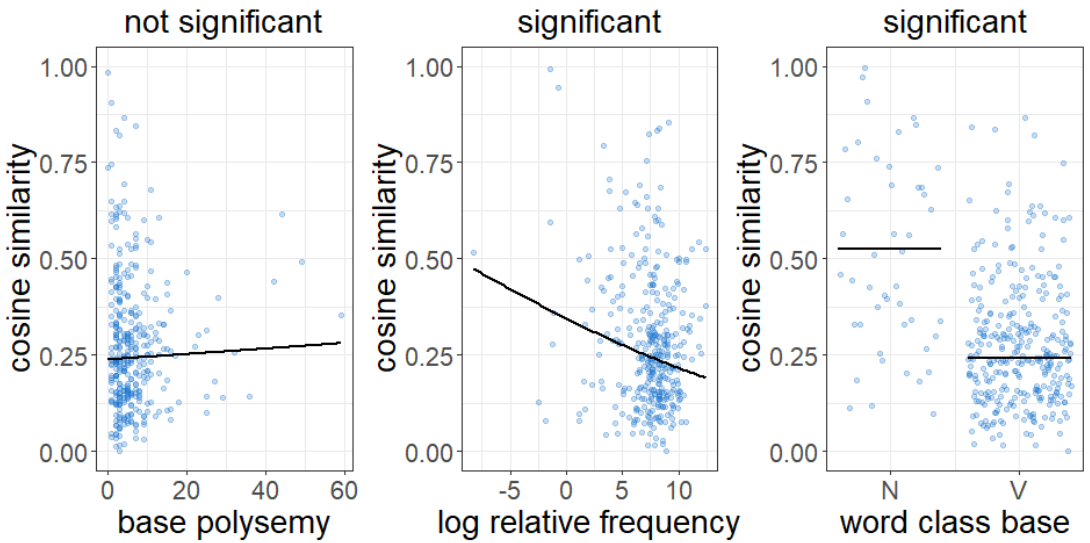
Similarity of nominal and verbal bases and derivatives for -*ee*

•Cosine similarity of denominal derivatives and nominal bases higher than that of deverbal derivatives and verbal bases
•Contra expectation that deverbal derivatives more similar to verbal bases

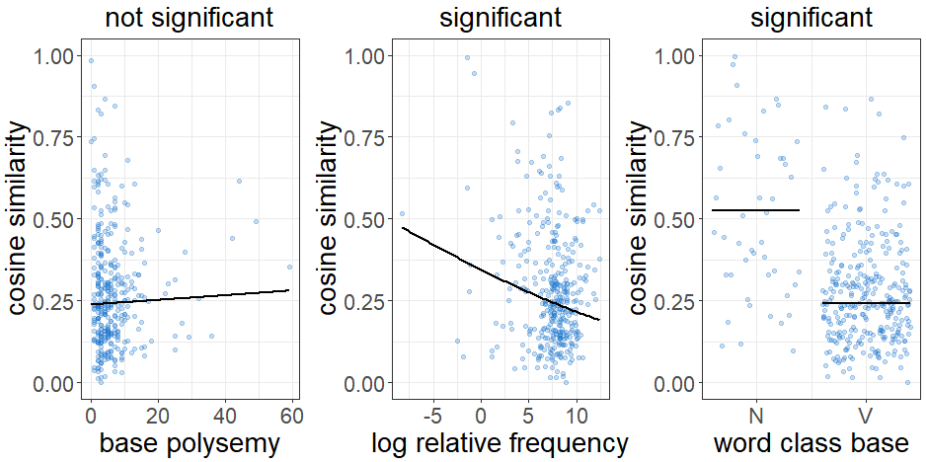


Beta regression model -*ee*

- Dependent variable cosine similarity

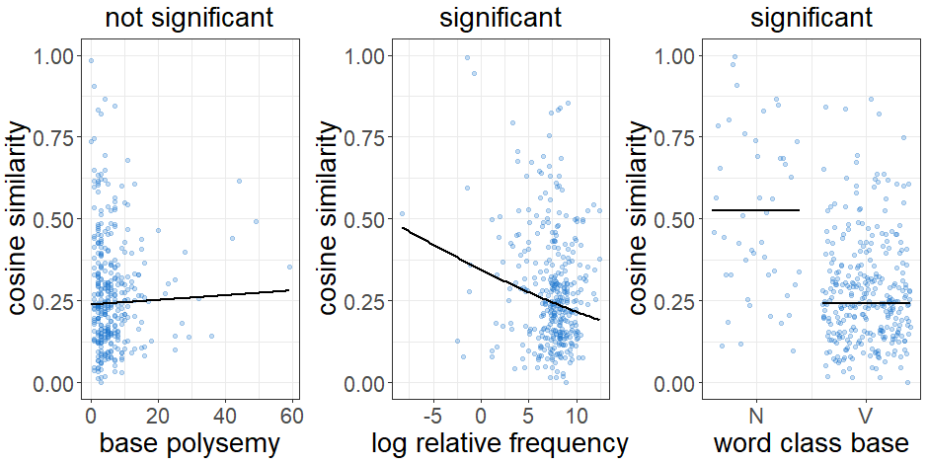


Beta regression model -*ee*



Beta regression model -*ee*

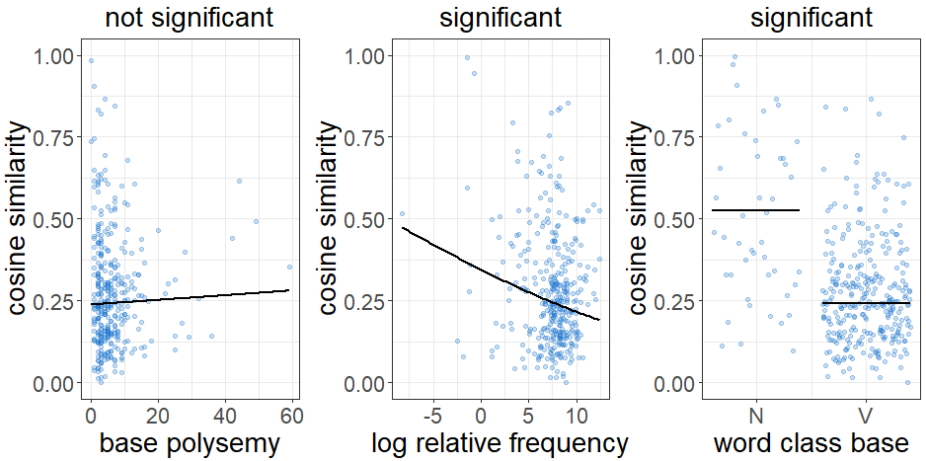
- Polysemy of base
  - Not significant



Beta regression model -*ee*

- Polysemy of base
  - Not significant

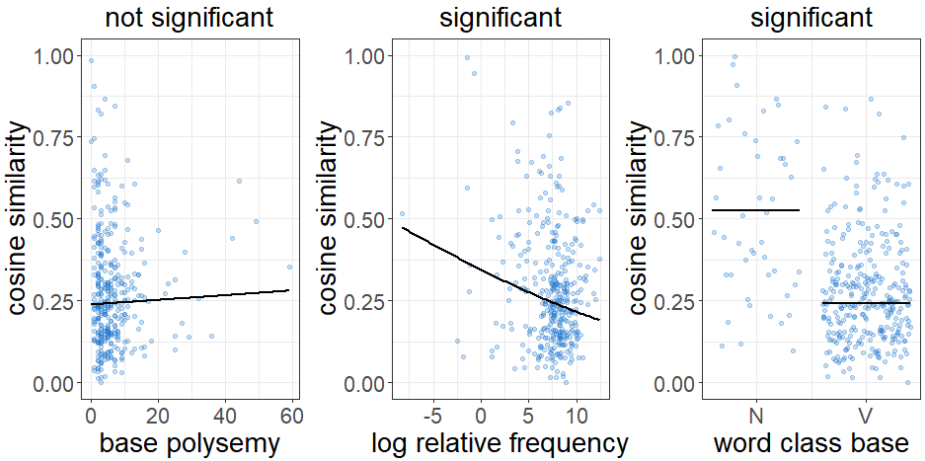
- Relative frequency
  - Significant
  - Higher relative frequency decreases cosine similarity
  - Not expected



Beta regression model -*ee*

- Polysemy of base
  - Not significant

- Relative frequency
  - Significant
  - Higher relative frequency decreases cosine similarity
  - Not expected



- Word class base
  - Significant
  - Cosine similarity decreases if base is a verb
  - Not expected

concertation

pixelation

The suffix *-ation*

67 denominal 72 deverbal

instrumentation

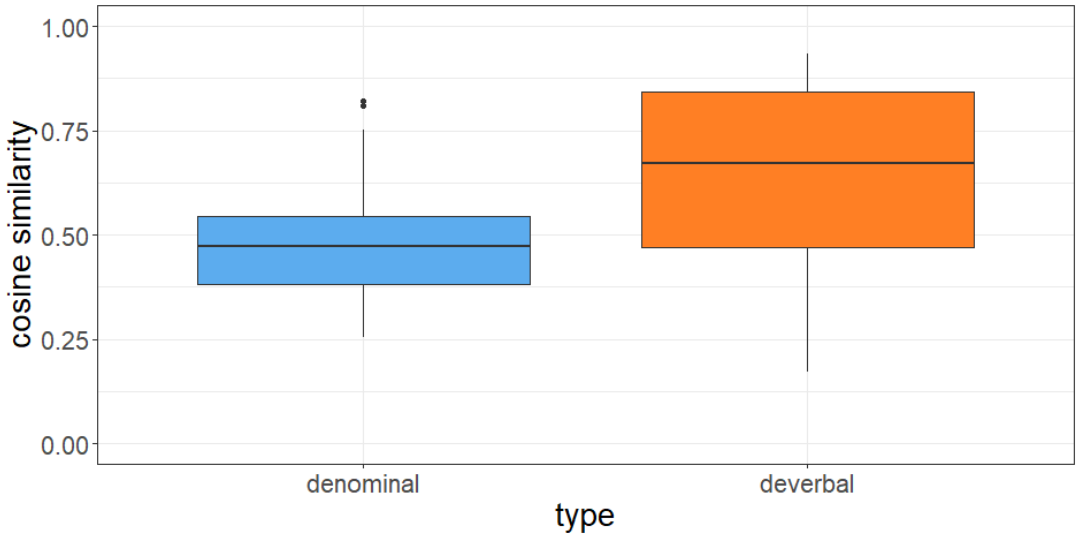
ozonation

impactation

Data from COCA (Davies 2008) and BNC (Davies 2004)

Similarity of nominal and verbal bases and derivatives for -*ation*

- Denominal derivatives and nominal bases show lower cosine similarity than deverbal pairs → opposite picture than for -*ee*



Beta regression model – Principal component analysis (PCA)

Beta regression model – Principal component analysis (PCA)

- Problem
  - Correlations of relative frequency, base polysemy, word class
  - Collinearity in model

Beta regression model – Principal component analysis (PCA)

- Problem
  - Correlations of relative frequency, base polysemy, word class
  - Collinearity in model

- Solution applied: PCA
  - Dimensionality of data reduced by transformation of problematic variables into principal components
  - Transformations lead to linear combinations of predictors
  - Resulting principal components are not correlated

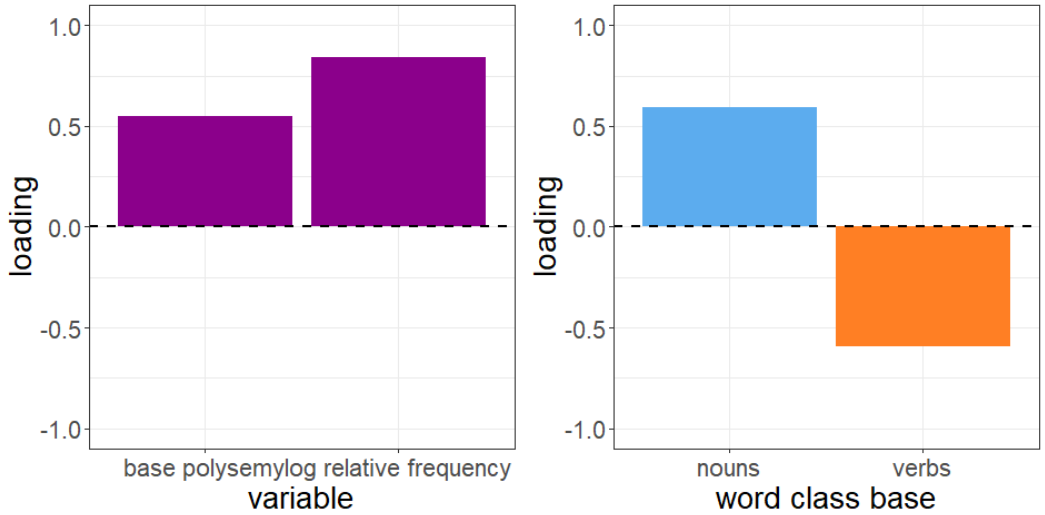
Beta regression model – Principal component analysis (PCA)

- Problem
  - Correlations of relative frequency, base polysemy, word class
  - Collinearity in model

- Solution applied: PCA
  - Dimensionality of data reduced by transformation of problematic variables into principal components
  - Transformations lead to linear combinations of predictors
  - Resulting principal components are not correlated

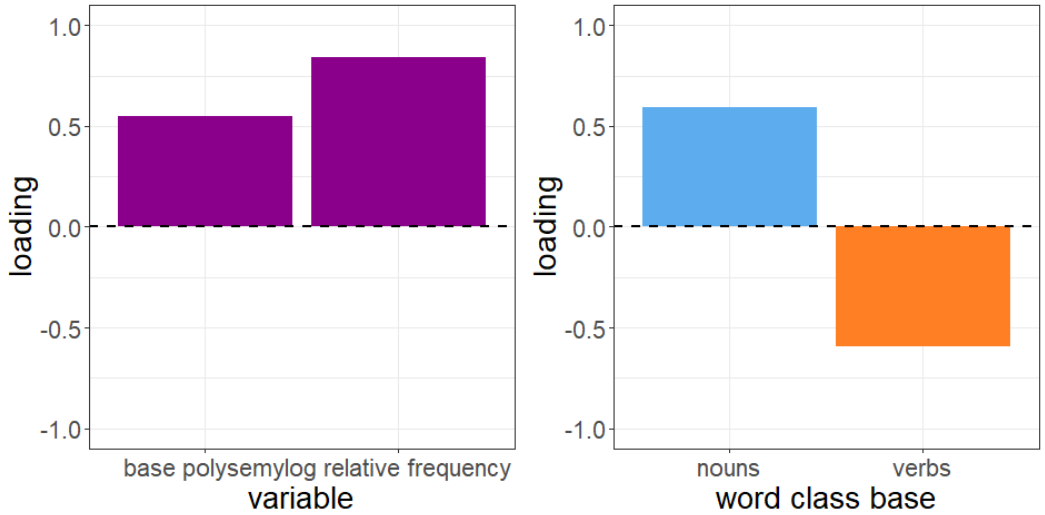
- First principal component is retained for analysis as fulfills common criteria (e.g., O'Rourke et al. 2005; Baayen 2008; Schmitz et al. 2021, Schmitz 2022)

Beta regression with PC for -*ation*



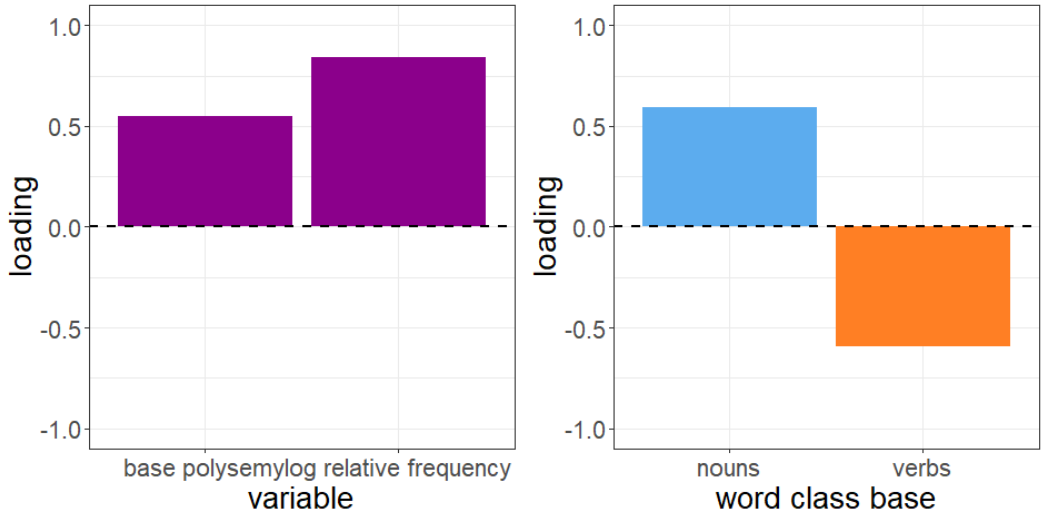
Beta regression with PC for -*ation*

- Loadings of retained PC



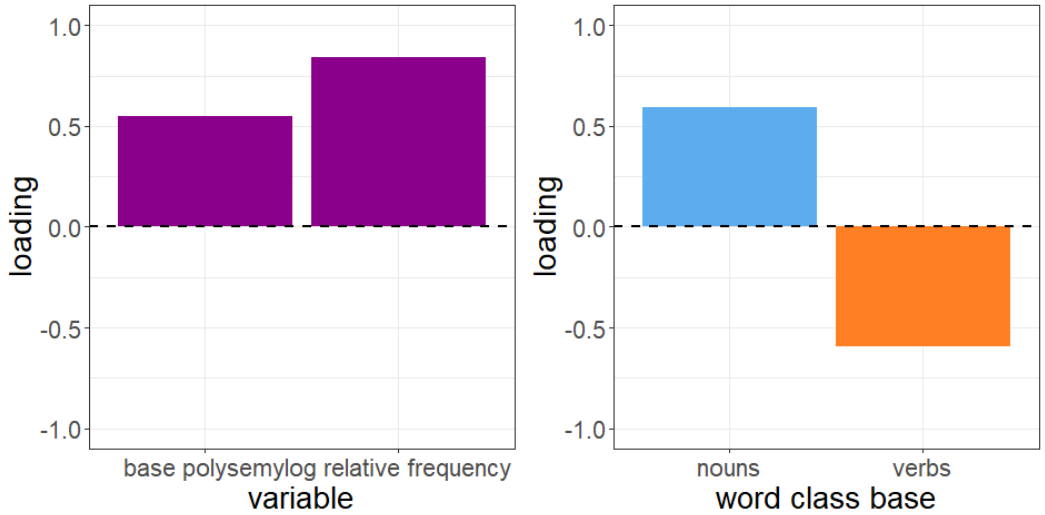
Beta regression with PC for -*ation*

- Loadings of retained PC
  - Polysemy of base and relative frequency same direction, relative frequency stronger represented in PC



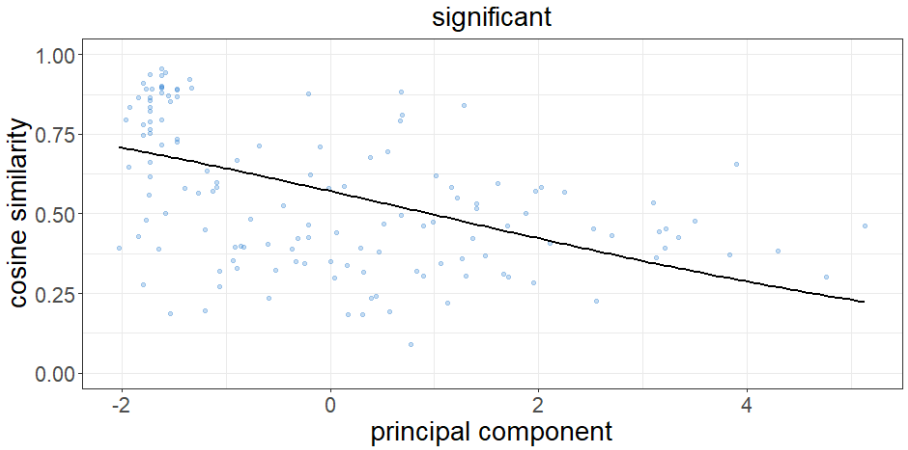
Beta regression with PC for -*ation*

- Loadings of retained PC
  - Polysemy of base and relative frequency same direction, relative frequency stronger represented in PC
  - Word classes point to different directions



Beta regression with PC for -*ation*

- Effect of retained PC
  - Higher polysemy of base word decreases cosine similarity (expected)
  - Higher relative frequency decreases cosine similarity (unexpected)
  - Word class of base influences cosine similarity (verbs higher cosine similarity, expected)

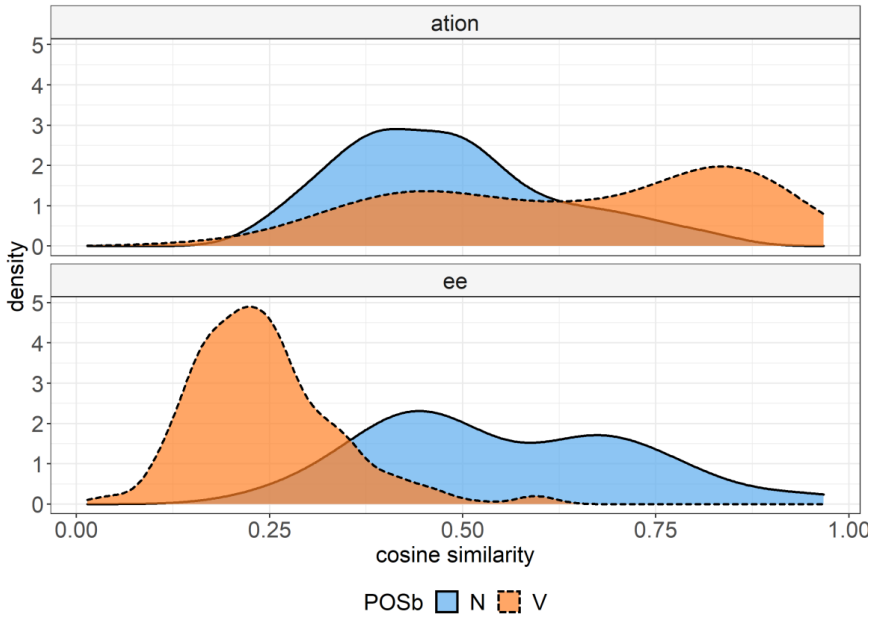


Summary -*ation* and -*ee*

Cosine similarity

Summary -*ation* and -*ee*

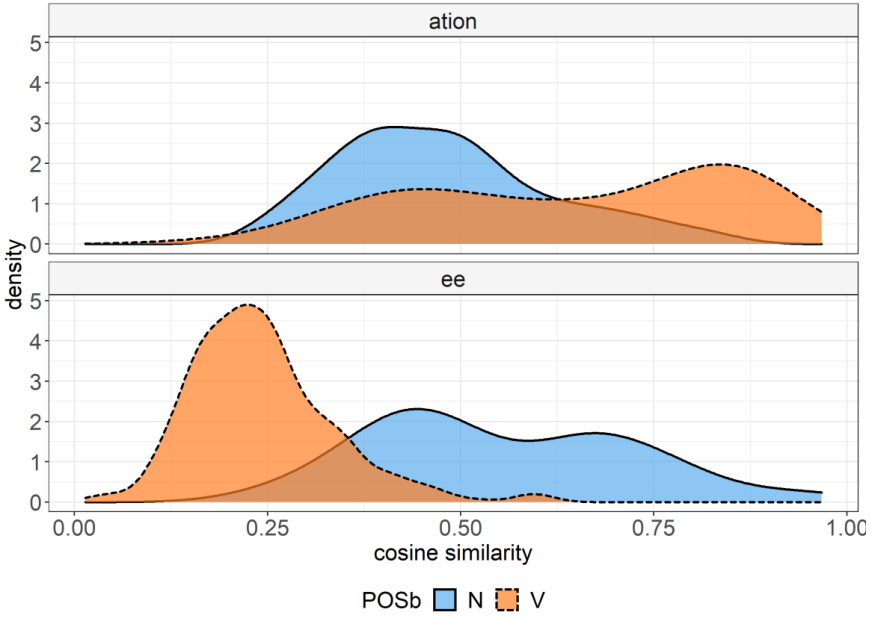
Cosine similarity



Summary -*ation* and -*ee*

Cosine similarity

- **Deverbal** -*ation* derivatives more similar to verbal bases than **denominal** derivatives to nominal bases

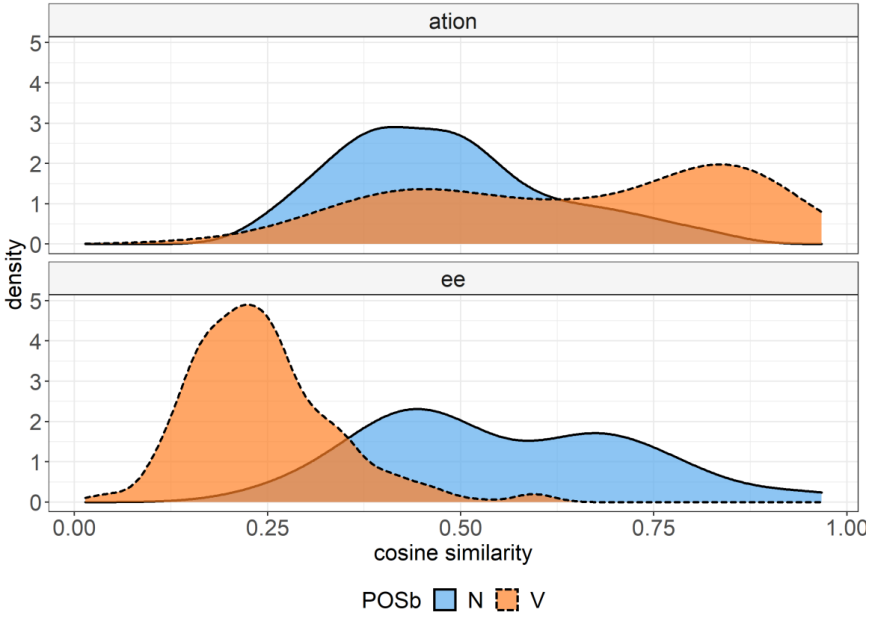


Summary -*ation* and -*ee*

Cosine similarity

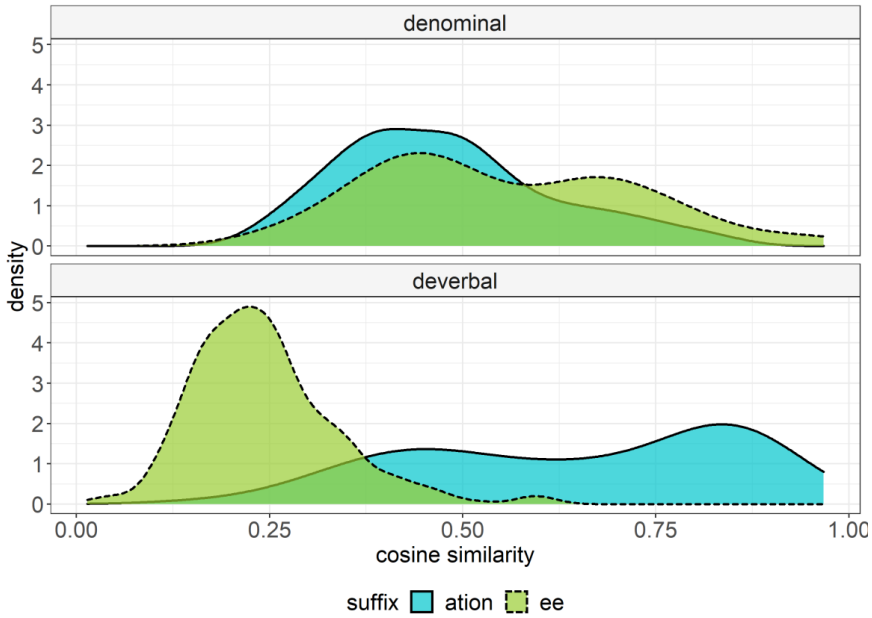
- **Deverbal** -*ation* derivatives more similar to verbal bases than **denominal** derivatives to nominal bases

- **Denominal** -*ee* derivatives more similar to nominal bases than **deverbal** derivatives to verbal bases



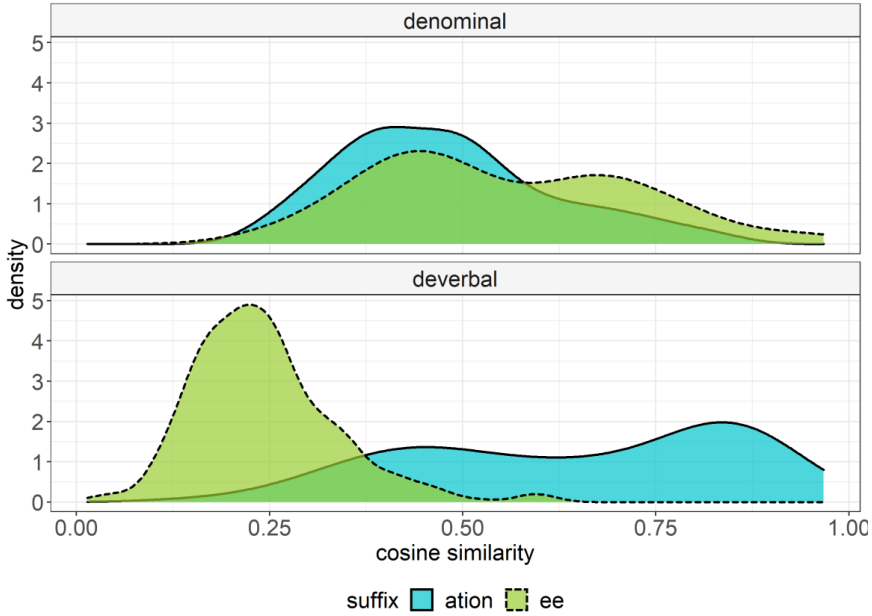
Differences denominal and deverbal

Differences denominal and deverbal



Differences denominal and deverbal

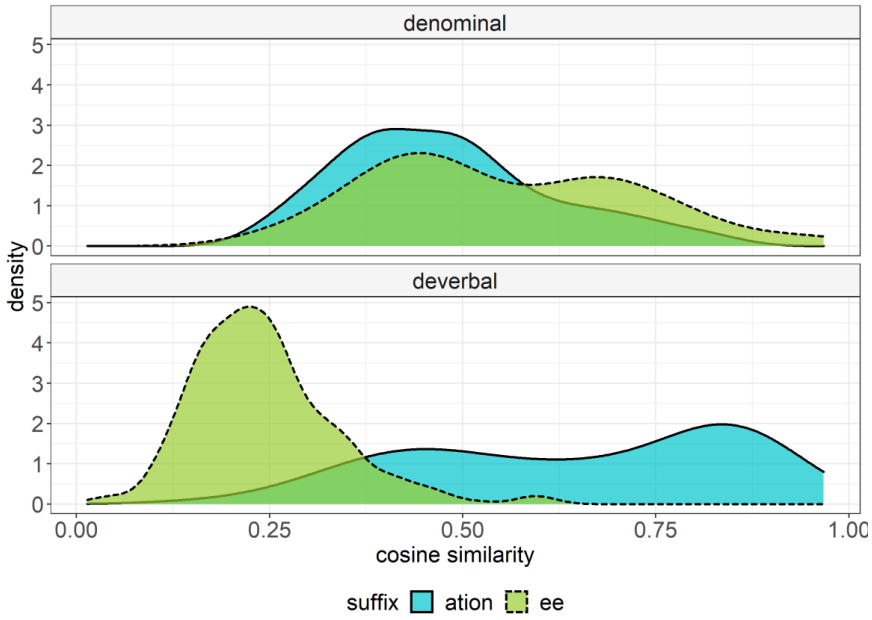
Denominal -*ation*
derivatives



Differences denominal and deverbal

Denominal -*ation* derivatives

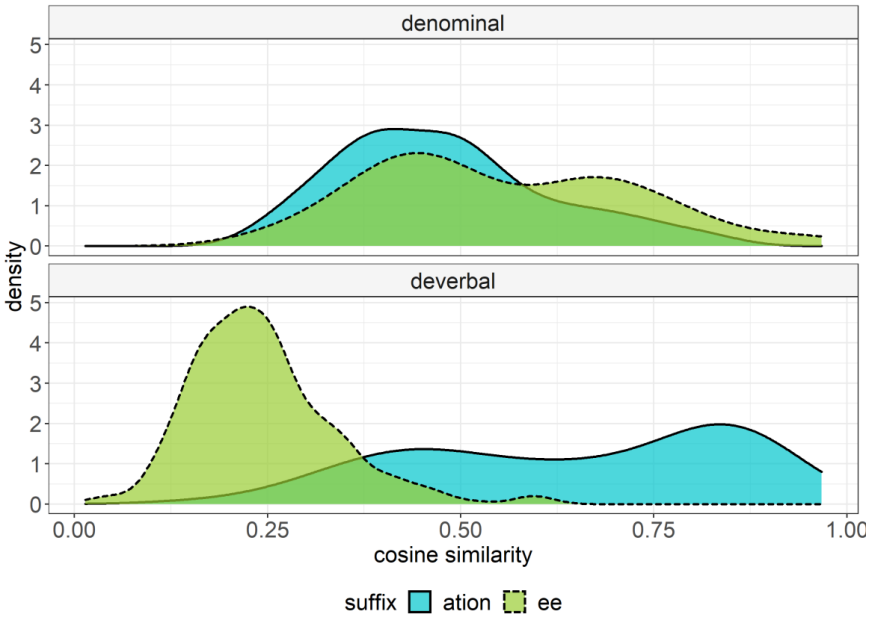
- are more uniform in their similarity than -*ee* derivatives



Differences denominal and deverbal

Denominal -*ation* derivatives

- are more uniform in their similarity than -*ee* derivatives
- overall less similar to their bases than -*ee* derivatives

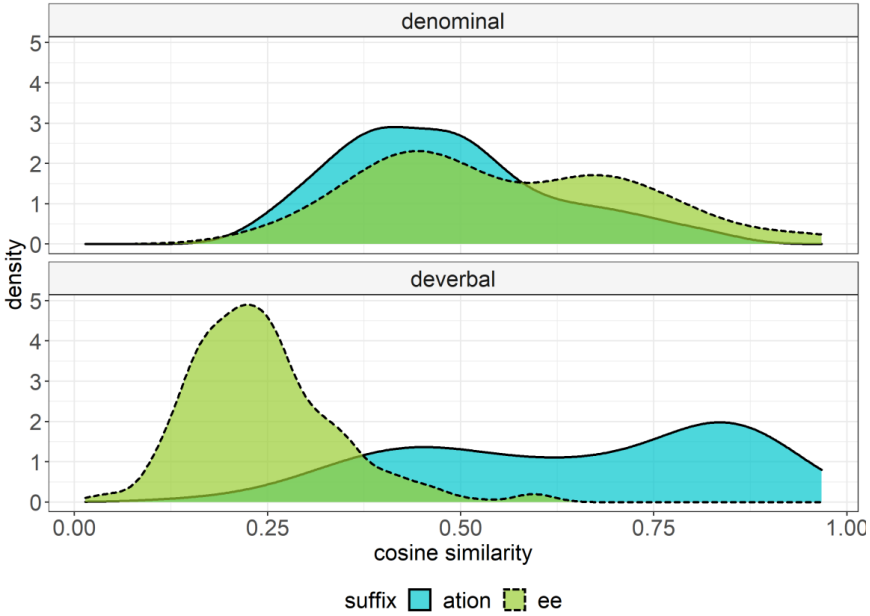


Differences denominal and deverbal

Denominal -*ation* derivatives

- are more uniform in their similarity than -*ee* derivatives
- overall less similar to their bases than -*ee* derivatives

Deverbal -*ee* derivatives



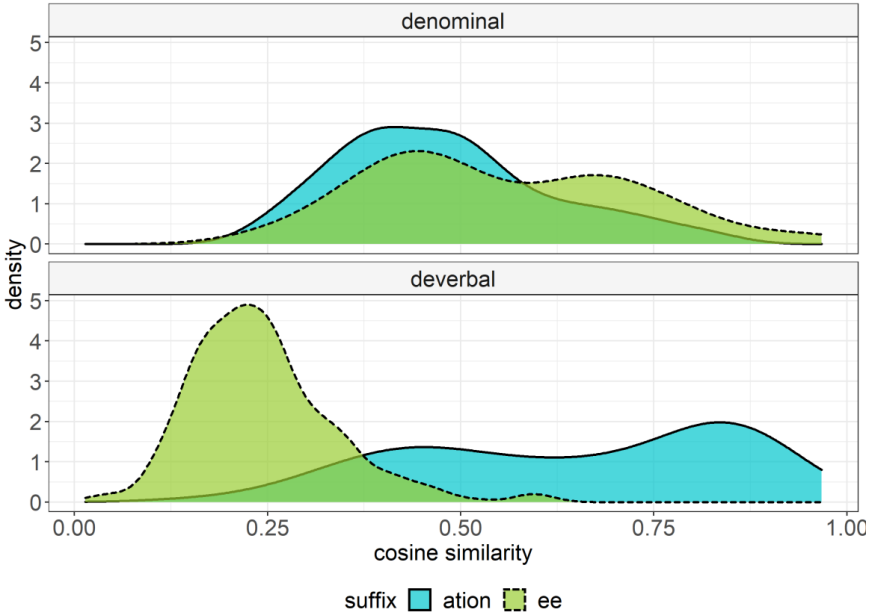
Differences denominal and deverbal

Denominal -*ation* derivatives

- are more uniform in their similarity than -*ee* derivatives
- overall less similar to their bases than -*ee* derivatives

Deverbal -*ee* derivatives

- are more uniform in their similarity than –*ation* derivatives



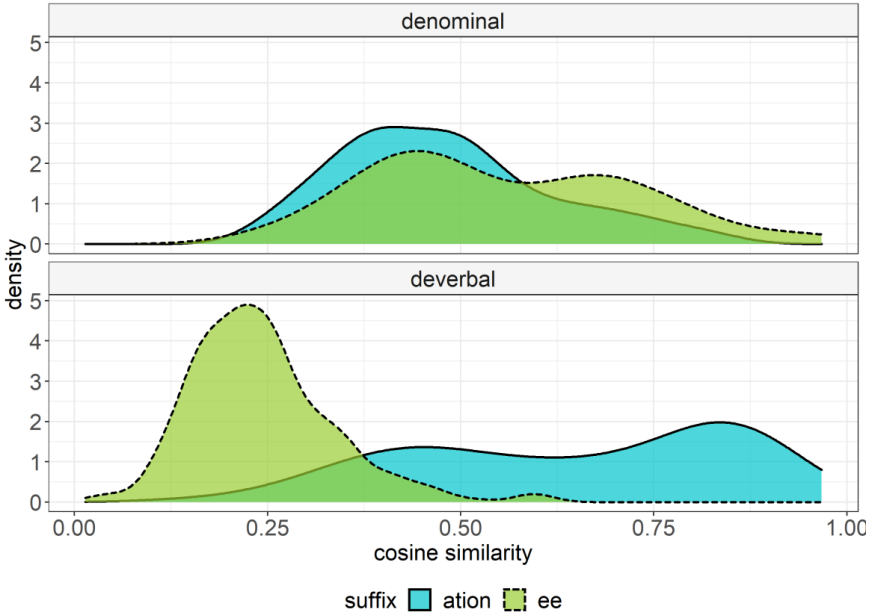
Differences denominal and deverbal

Denominal -*ation* derivatives

- are more uniform in their similarity than -*ee* derivatives
- overall less similar to their bases than -*ee* derivatives

Deverbal -*ee* derivatives

- are more uniform in their similarity than –*ation* derivatives
- overall much less similar to their bases than –*ation* derivatives



Discussion of effects

Discussion of effects

- Cosine similarity significantly influenced by

Discussion of effects

- Cosine similarity significantly influenced by
  - Relative frequency for both data sets (contra expectation)

Discussion of effects

- Cosine similarity significantly influenced by
  - Relative frequency for both data sets (contra expectation)
  - Polysemy of base for -*ation* data (in line with expectation)

Discussion of effects

- Cosine similarity significantly influenced by
  - Relative frequency for both data sets (contra expectation)
  - Polysemy of base for -*ation* data (in line with expectation)
  - Word class of base for both data sets (contra expectation for -*ee*, in line with expectation for -*ation*)

Discussion of effects

- Cosine similarity significantly influenced by
  - Relative frequency for both data sets (contra expectation)
  - Polysemy of base for -*ation* data (in line with expectation)
  - Word class of base for both data sets (contra expectation for -*ee*, in line with expectation for -*ation*)

- Difference in effect of word class of base due to ontology?
  (e.g., Van Valin & LaPolla 1997; Haspelmath 2001; Szabó 2015)

Discussion of effects

- Cosine similarity significantly influenced by
  - Relative frequency for both data sets (contra expectation)
  - Polysemy of base for -*ation* data (in line with expectation)
  - Word class of base for both data sets (contra expectation for -*ee*, in line with expectation for -*ation*)

- Difference in effect of word class of base due to ontology?
  (e.g., Van Valin & LaPolla 1997; Haspelmath 2001; Szabó 2015)
    - -*ee* creates participant readings → participants usually denoted by nouns

Discussion of effects

- Cosine similarity significantly influenced by
  - Relative frequency for both data sets (contra expectation)
  - Polysemy of base for -*ation* data (in line with expectation)
  - Word class of base for both data sets (contra expectation for -*ee*, in line with expectation for -*ation*)

- Difference in effect of word class of base due to ontology?
  (e.g., Van Valin & LaPolla 1997; Haspelmath 2001; Szabó 2015)
  - -*ee* creates participant readings → participants usually denoted by nouns
  - -*ation* refers to eventualities → eventualities usually denoted by verbs

Open questions & future directions

Open questions & future directions

- Is there also a difference for nominalizations with other suffixes, e.g., -*ment*?

Open questions & future directions

- Is there also a difference for nominalizations with other suffixes, e.g., -*ment*?

- Do we find similar results with a different methodology, e.g., Linear Discriminative Learning (LDL)?
  - Error-driven learning → different vectors
  - Different measures → explanation for findings?

References

- Alexiadou, Artemis. 2010. Nominalizations: A probe into the architecture of grammar part i: The nominalization puzzle. *Language and Linguistics Compass* 4(7). 496-511.
- Baayen, R. Harald. (2008). *Analyzing linguistic data: A practical introduction to statistics using R*. Cambridge University Press. doi:10.1017/CBO9780511801686.
- Barker, Chris. 1998. Episodic -ee in English: A thematic role constraint on new word formation. *Language* 74 (4), pp. 695-727.
- Bauer, Laurie, Rochelle Lieber & Ingo Plag. 2013. *The Oxford reference guide to English morphology*. Oxford: Oxford University Press.
- Bojanowski, Piotr, Edouard Grave, Armand Joulin & Tomas Mikolov. 2016. Enriching word vectors with subword information. arXiv preprint arXiv:1607.04606 .
- Davies, Mark. 2004. *British National Corpus* (from Oxford University Press). Available online at https://www.english-corpora.org/bnc/.
- Davies, Mark. 2008. *The Corpus of Contemporary American English: 400+ million words, 1990-present*. Available online at https://www.english-corpora.org/coca/.
- Haspelmath, Martin. 2001. Word classes and parts of speech. In Neil J. Smelser & Paul B. Baltes (eds.), *International encyclopedia of the social & behavioral sciences*, 16538-16545. Amsterdam: Elsevier.
- Hay, Jennifer & Harald Baayen. 2003. Phonotactics, parsing and productivity. *Italian Journal of Linguistics* 1. 99–130. doi:10.1.1.171.705.
- Kawaletz, Lea. 2021. *The semantics of english -ment nominalizations*. PhD Dissertation, Heinrich Heine-Universität Düsseldorf.
- Kawaletz, Lea & Ingo Plag. 2015. Predicting the semantics of English nominalizations: A frame-based analysis of –ment Suffixation. In: *Semantics of complex words*. Bauer, Laurie, Lívia Körtvélyessy, Pavol Štekauer (Eds.), pp. 289-319.
- Lapesa, Gabriella, Lea Kawaletz, Ingo Plag, Marios Andreou, Max Kisselew & Sebastian Padó. 2018. Disambiguation of newly derived nominalizations in context: A distributional semantics approach. *Word Structure* 11(3). 277–312. doi: 10.3366/word.2018.0131.
- Mikolov, Tomas, Edouard Grave, Piotr Bojanowski, Christian Puhrsch & Armand Joulin. 2018. *Advances in Pre-Training Distributed Word Representations*.
- O'Rourke, Norm; Hatcher, Larry; Stepanski, E. J. (2005). Using SAS for Univariete & Multivariate Statistics. SAS Institute Inc.
- Plag, Ingo. 1999. *Morphological productivity: Structural constraints in English derivation*. Berlin: Mouton de Gruyter.
- Plag, Ingo. 2004. Syntactic category information and the semantics of derivational morphological rules. *Folia Linguistica* 38(3-4). 193–225.
- Plag, Ingo, Marios Andreou & Lea Kawaletz. 2018. A frame-semantic approach to polysemy in affixation. In Olivier Bonami, Gilles Boyé, Georgette Dal, Hélène Giraudo & Fiammetta Namer (eds.), *The lexeme in descriptive and theoretical morphology*, 467–486. Berlin: Language Science Press.
- Schmitz, Dominic. 2022. Production, perception, and comprehension of subphonemic detail. Language Science Press. https://langsci-press.org/catalog/book/365.
- Schmitz, D., Plag, I., Baer-Henney, D., & Stein, S. D. (2021). Durational Differences of Word-Final /s/ Emerge From the Lexicon: Modelling Morpho-Phonetic Effects in Pseudowords With Linear Discriminative Learning. Frontiers in Psychology, 12, 2983.
- Schneider, Viktoria. 2023. Eventualities in the semantics of denominal nominalizations. In Sven Kotowski & Ingo Plag (eds.), The semantics of derivational morphology: Theory, methods, evidence, de Gruyter. In press.
- Szabó, Zoltán Gendler. 2015. Major parts of speech. *Erkenntnis* 80(S1). 3-29.
- Van Valin, Robert D. & Randy J. LaPolla. 1997. *Syntax: Structure, meaning and function*. Cambridge textbooks in linguistics. Cambridge: Cambridge Univ. Press.

hhu Heinrich Heine Universität Düsseldorf

Thank you!

Method: Data

- Corpus search on BNC (Davies 2004) and COCA (Davies 2008) for nominalizations with suffixes –*ee* and -*ation*

- Identification of bases manually
  - *biography* → *biographee*
  - *ozone* → *ozonation*

- Determination of word class of base on basis of a frequency criterion
  - As soon as the verbal base represents over 30% of all tokens of the base forms → base deverbal

charge	Tokens	Percentage
V	8265	31
N	26469	69

Beta regression with principal component for -*ation*

Beta regression with principal component for *-ation*

- Common criteria PCA

Beta regression with principal component for *-ation*

- Common criteria PCA
  - Eigenvalue higher than 1

Beta regression with principal component for *-ation*

- Common criteria PCA
  - Eigenvalue higher than 1
  - Cumulative percentage explained higher than 80%

Beta regression with principal component for *-ation*

- Common criteria PCA
  - Eigenvalue higher than 1
  - Cumulative percentage explained higher than 80%
  - PC has to make sense in their loadings

Beta regression with principal component for *-ation*

- Common criteria PCA
  - Eigenvalue higher than 1
  - Cumulative percentage explained higher than 80%
  - PC has to make sense in their loadings
    - Here it decreases cosine similarity