

Han mannen and his neighbours:
A contrastive study of *han mannen's*
constructional neighbourhood in
Bokmål, Danish, and Swedish

BARTHE BLOOM

UNIVERSITY OF JENA

CONSTRUCTIONS IN THE NORDICS 2, NOVEMBER 16TH 2020, ONLINE

Overview

- Introduction
 - *Han mannen*
 - Background
 - Aim of the study
- Methodology
 - Data collection
 - Statistical methods
 - Variables
- Preliminary results
- Conclusion

Introduction

Han mannen

- 1) *Da tenker jeg nok at **han mannen** i butikken har helt rett.*
(NoTenTen 17, Bokmål)

Colloquial

Background deixis (Lie 2010a, 2010b)

Psychological distance (Johannessen 2006, 2008, 2020)

- Pragmatic inference
- *Hun jenta, han dansken, han læreren, han treåringen*

Han mannen

1) *Da tenker jeg nok at **han mannen** i butikken har helt rett.*

(NoTenTen 17, Bokmål)

2) *Om detta vill inte Fredrik statsministern prata, **han mannen** som de arbetslösa ungdomarna i Eskiltuna tyckte var kall efter att ha träffat honom.*

(SvTenTen 14, Swedish)

3) *Men som Fluesmækkeren sikkert siger, så er der penge i lortet. Hvorfor hedder **ham manden** Fluesmækkeren?*

(DaTenTen 14, Danish)

Today's micro-constructions

Bokmål & Swedish

[*han*_{DET} *mann*_{SG} *-en*_{ART}]

[*mann*_{SG} *-en*_{ART}]

[*den*_{DET} *mann*_{SG} *-en*_{ART}]

Danish

[*ham*_{DET} *mand*_{SG} *-en*_{ART}]

[*mand*_{SG} *-en*_{ART}]

[*den*_{DET} *mand*_{SG}]

Aim of the current study

To gain insight into the differences and similarity in the make-up of constructional neighborhoods of closely related languages by investigating *han mannen*'s sequential relations and how these relate to *den mannen* and *mannen*'s sequential relations in Bokmål, Danish, and Swedish.

Constructional neighborhoods

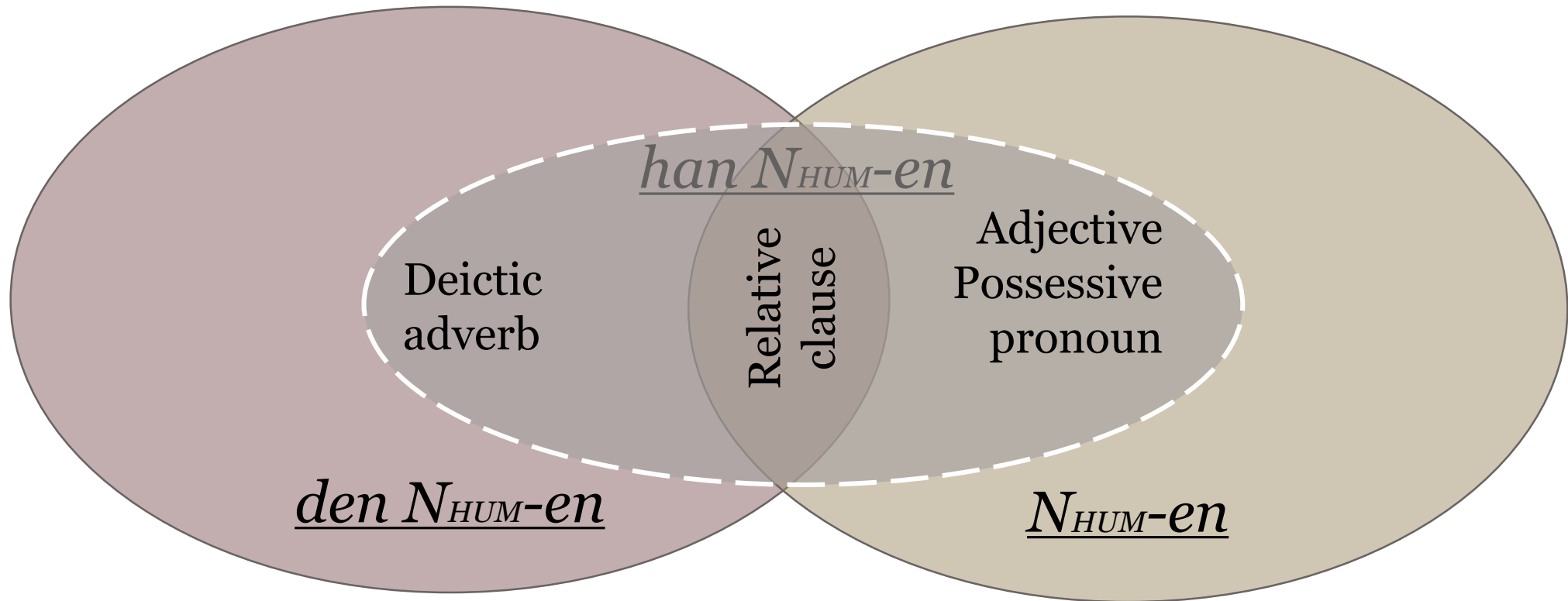
Nested-network approach (Diessel 2019)

Constructions at the same level of abstraction are connected

Relational similarity and contrast

Location in the network (Diessel 2019, 200)

Overlap in sequential relations (Spoken Norwegian)



Methodology

Data collection

Informal written data

Corpora

- Bokmål: *noTenTen17, Bokmål* (2 903 629 903 tokens)
- Danish: *daTenTen17* (2 566 932 492 tokens)
- Swedish: *svTenTen14* (3 889 895 434 tokens)

Data collection

Concordance search: e.g., “han mannen” and “han * mannen”

First clean up (e.g., two distinct NPs, other languages)

	Unfiltered	Post clean up
Bokmål	3182	1176
Danish	780	343
Swedish	1916	63

Data set

	Bokmål	Danish	Swedish	total
<i>den mannen</i>	75	75	75	225
<i>han mannen</i>	75	75	63	213
<i>mannen</i>	75	75	75	225
total	225	225	213	663

Statistical methods

Conditional inference trees (Hothorn et al. 2006)

- Optimal classification of the data

Random forests (Strobl et al. 2008)

- Compute a multitude of conditional inference trees, but based on random samples of the data and the variables (Tagliamonte and Baayen 2012).

Partial dependence scores (Greenwell 2017)

- Represent the direction of the effect that the values of the predictor variables have on the response variable (Gries 2019).

Variables

Response:

- Language bm/da/sv

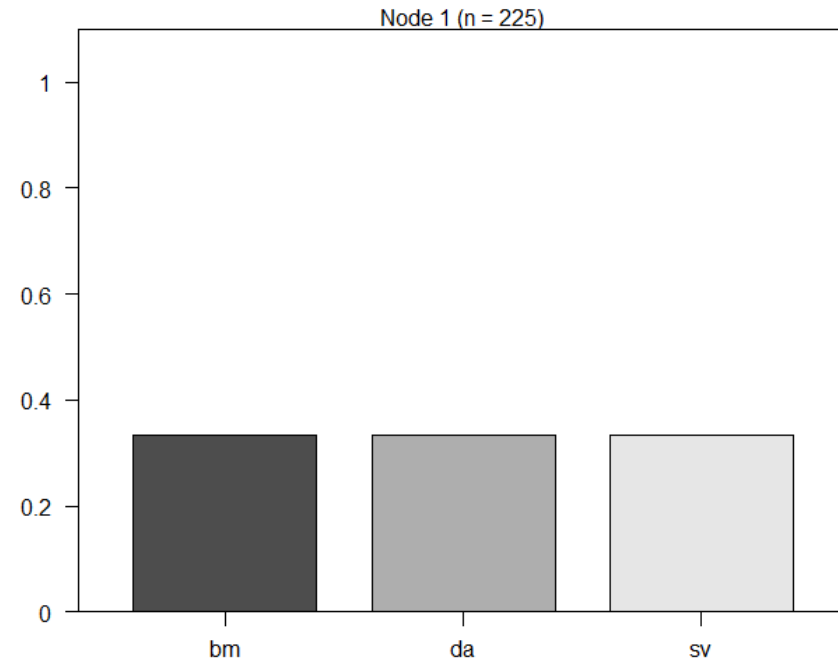
Predictors:

- Prenominal modification (mod/no)
- Deictic intensifying adverb (int/no)
- Possessive pronoun (poss/no)
- Prepositional phrase (pp/no)
- Relative clause (rc/no)

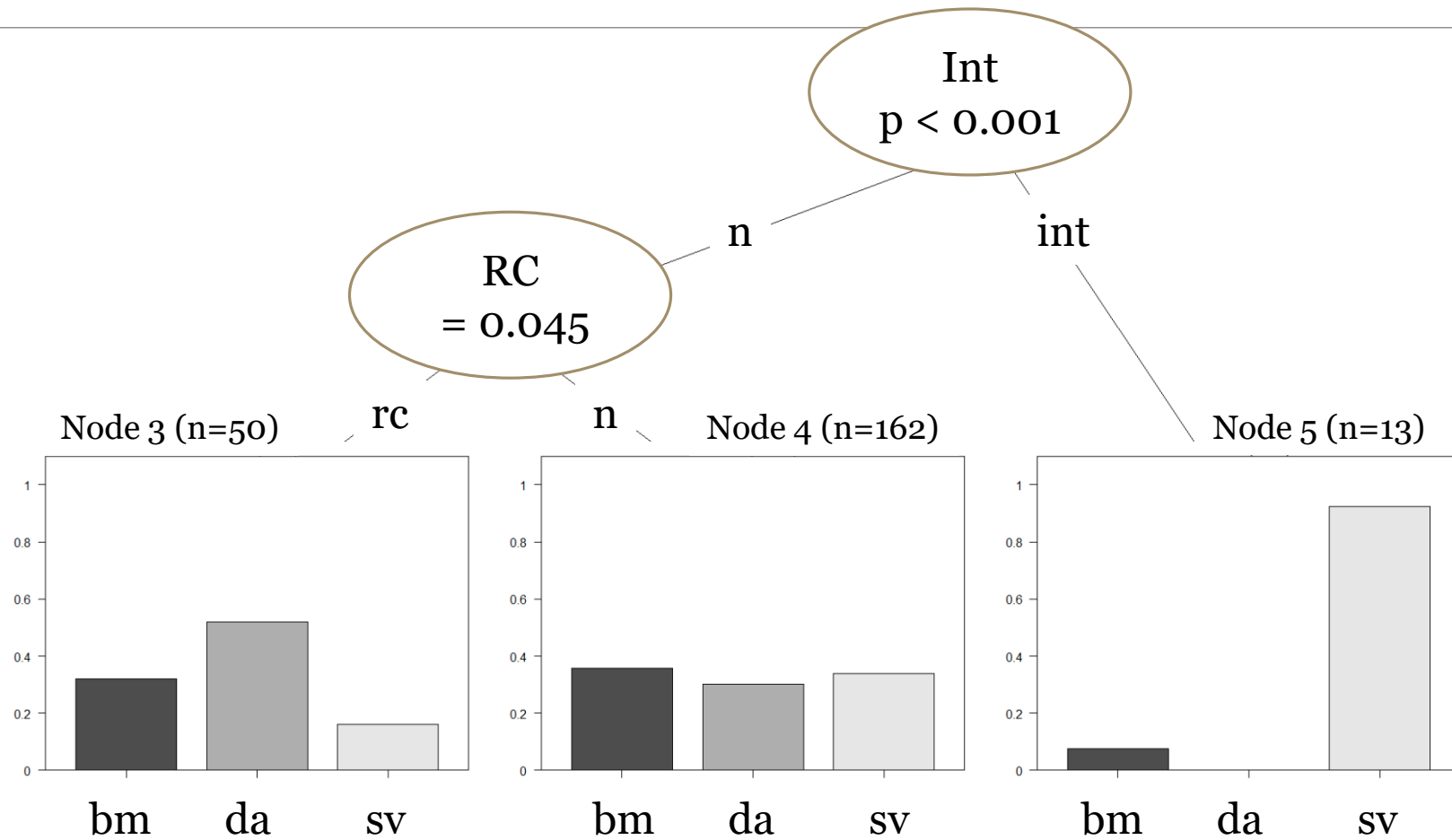
Preliminary results

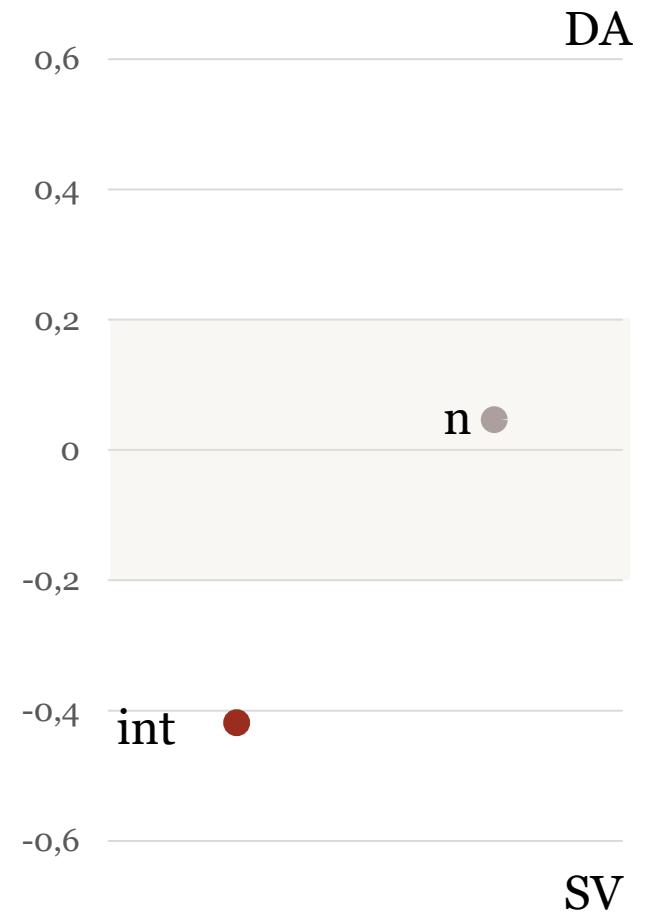
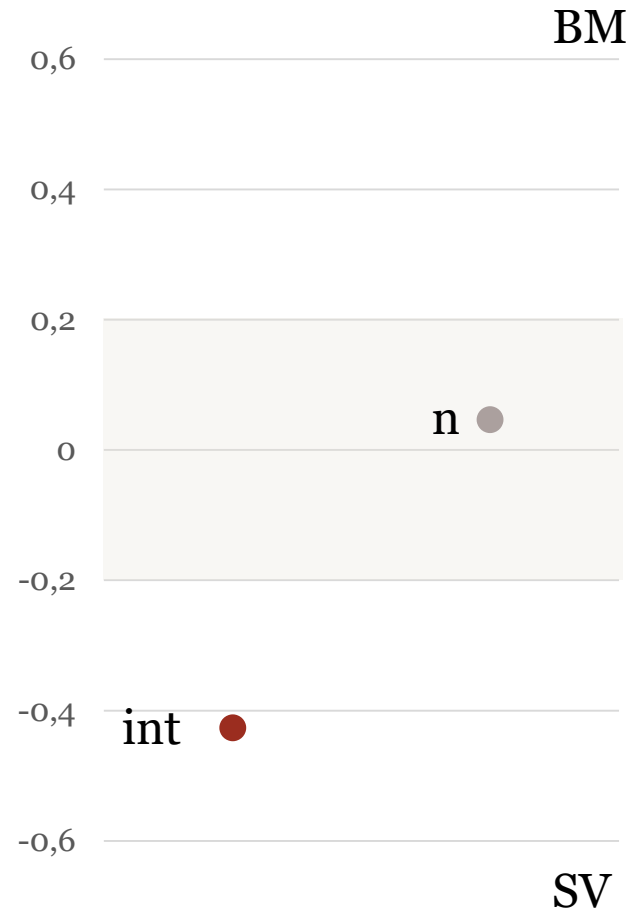
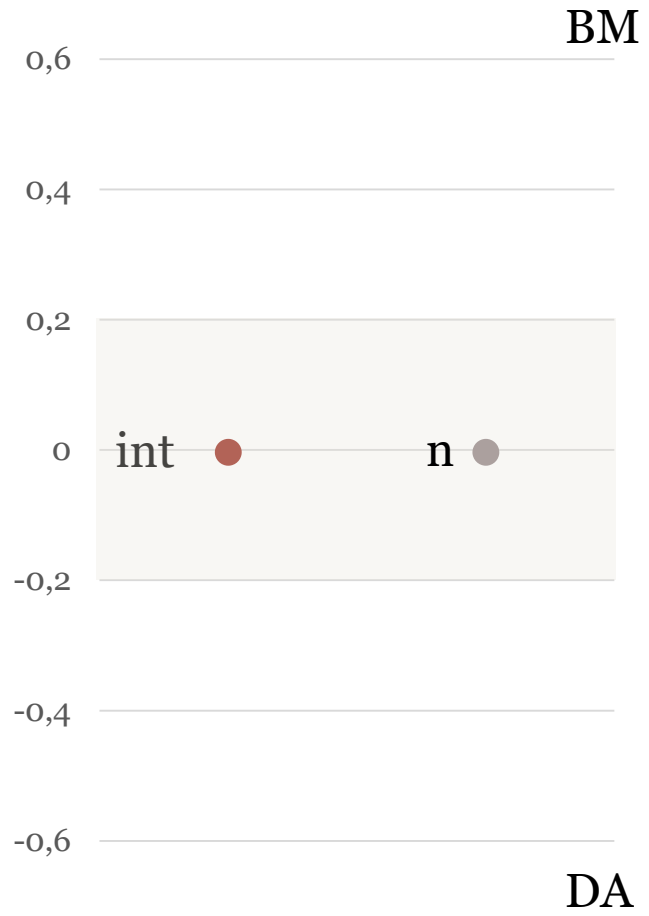
Mannen

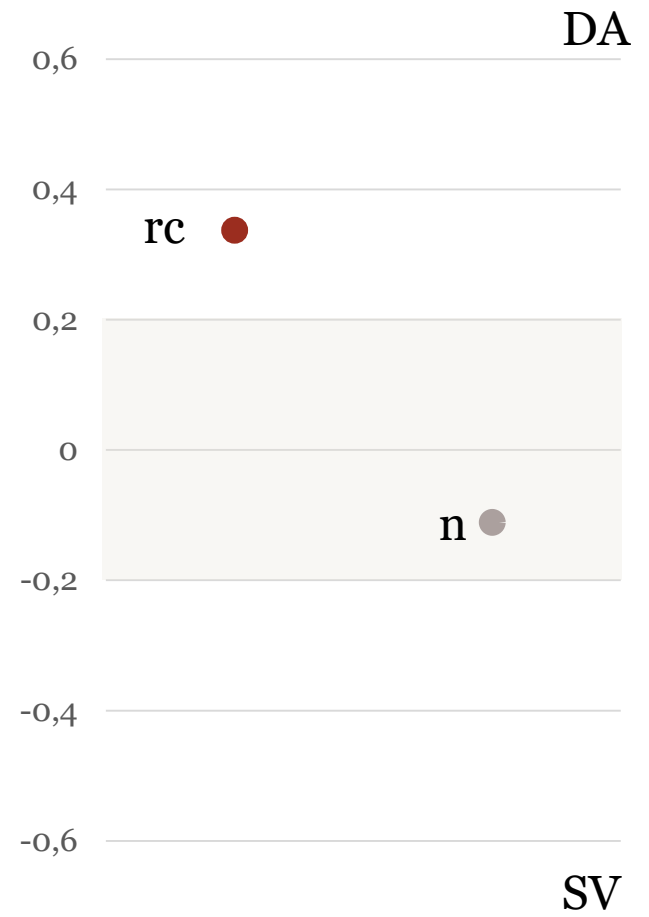
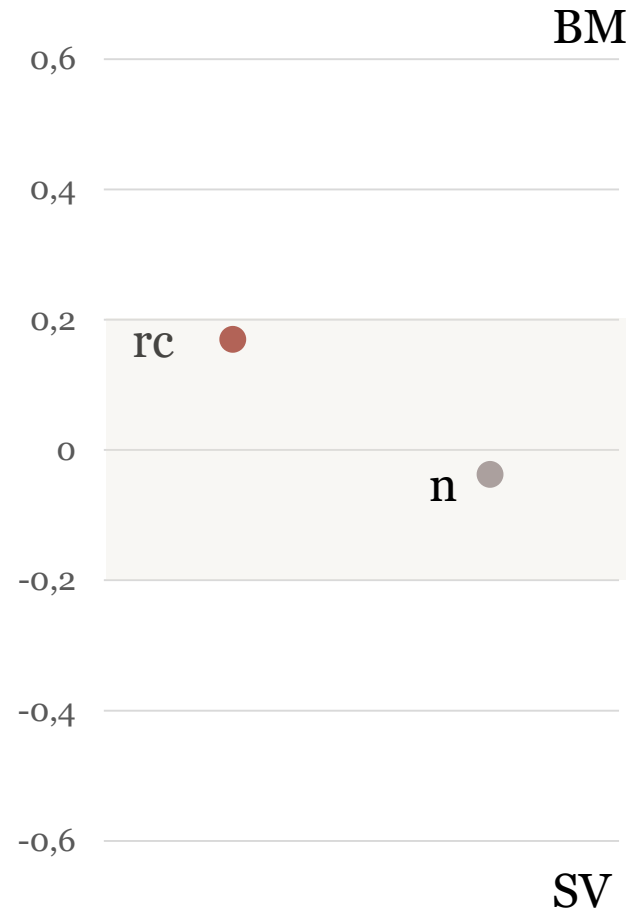
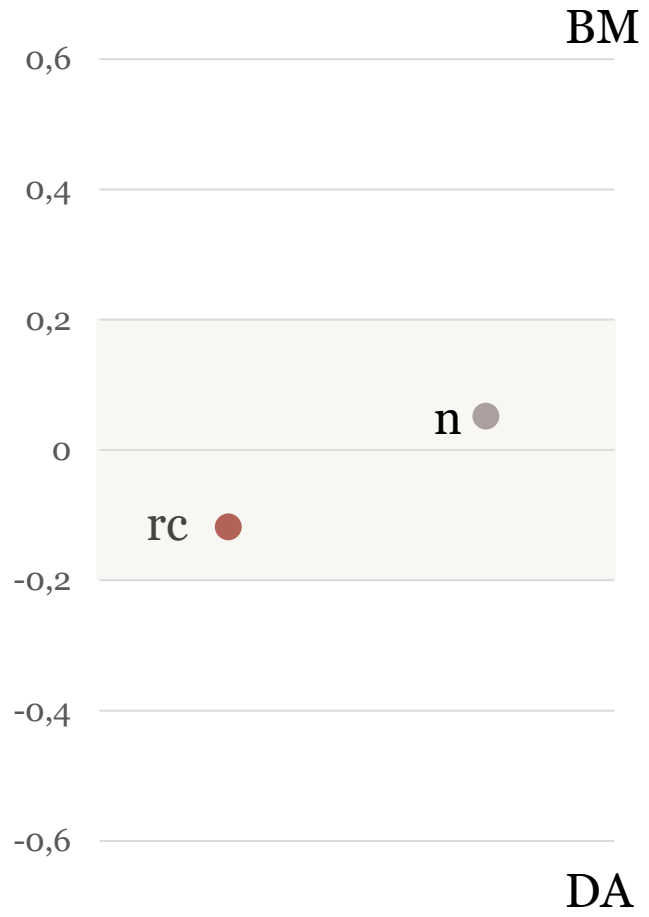
No distinctive sequential relations (at this point)



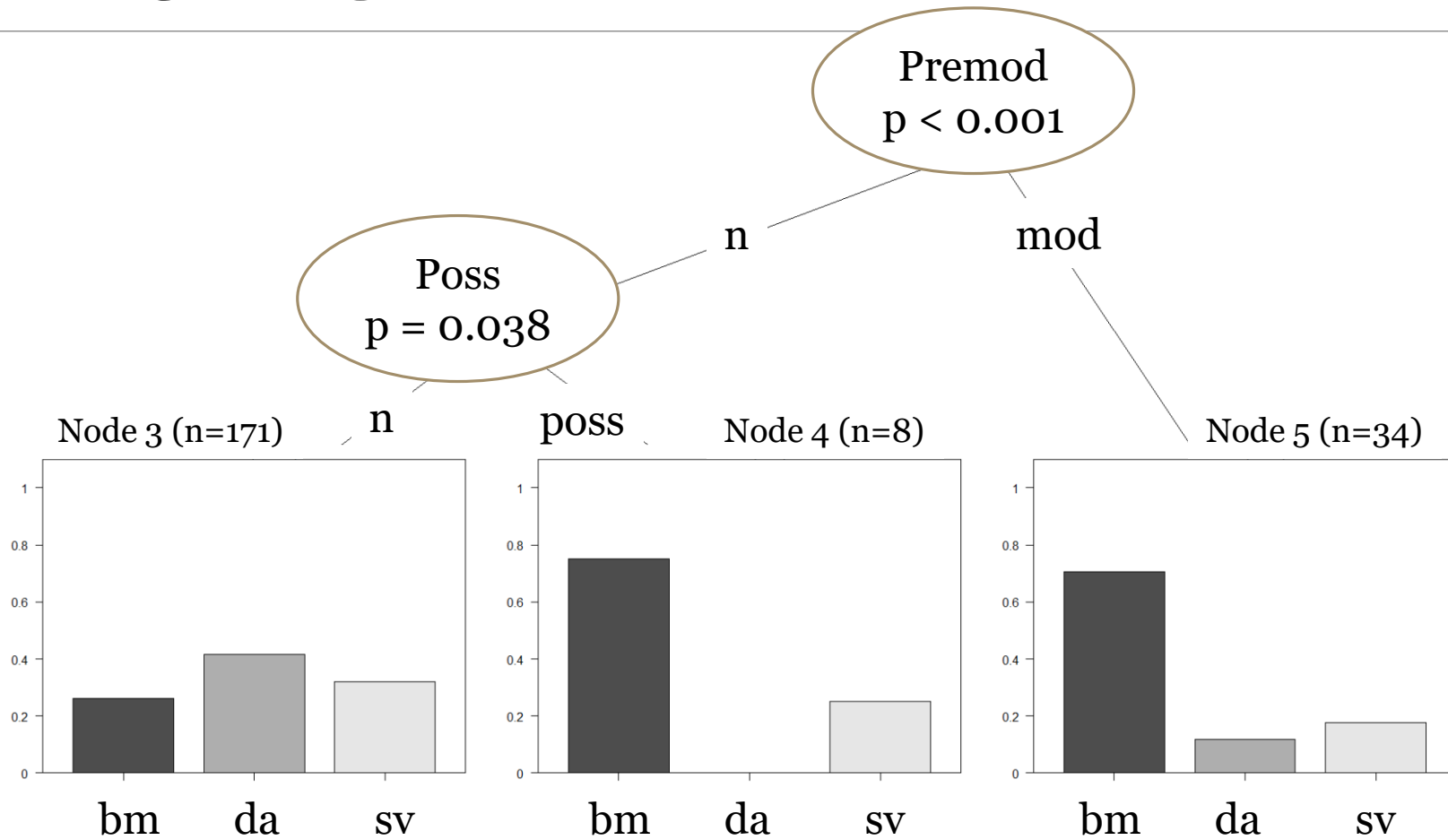
Den mannen

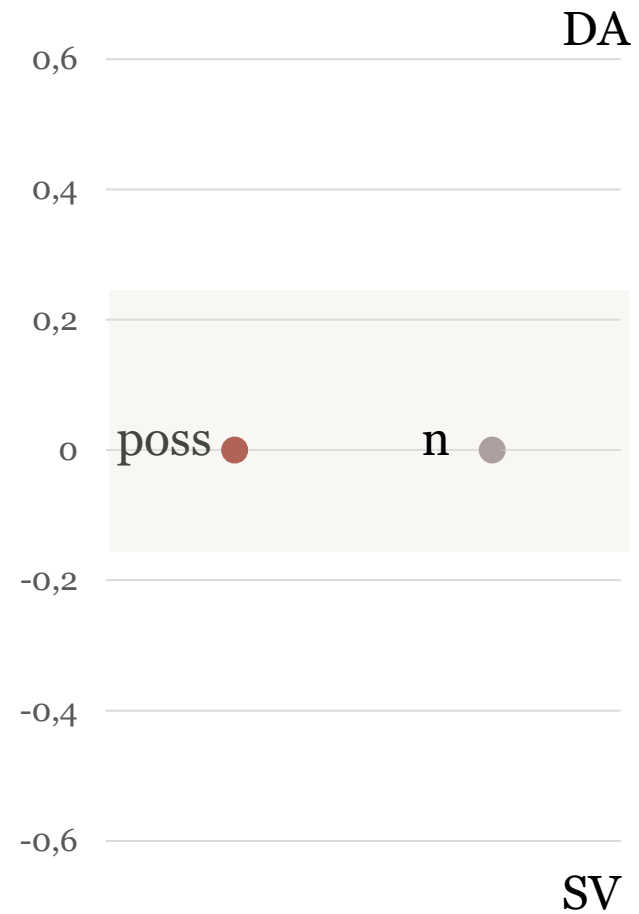
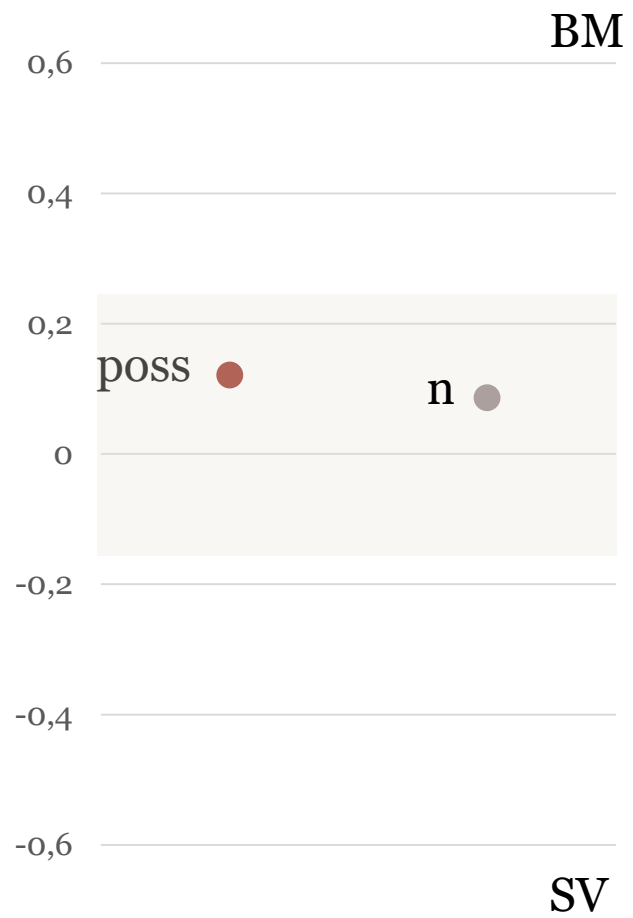
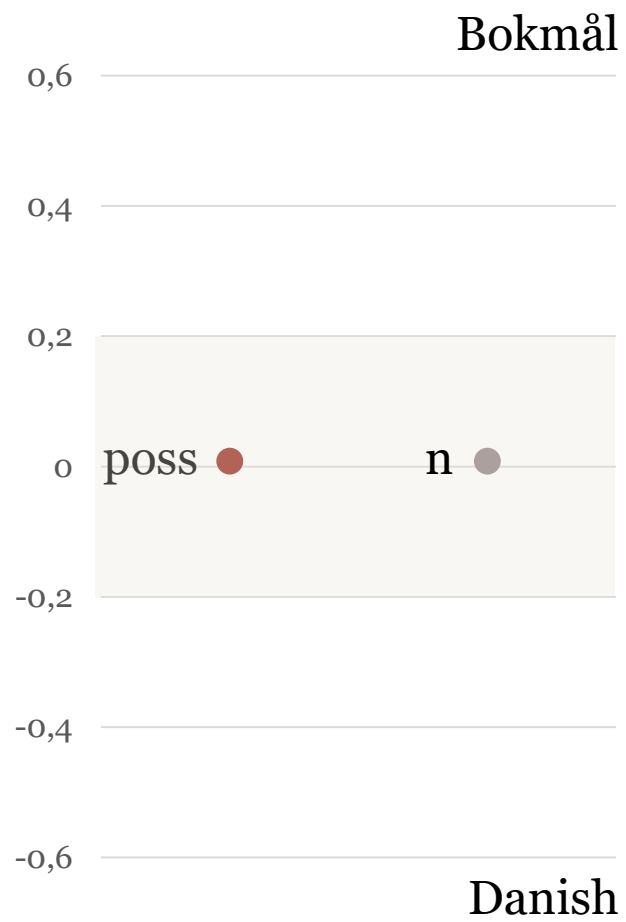


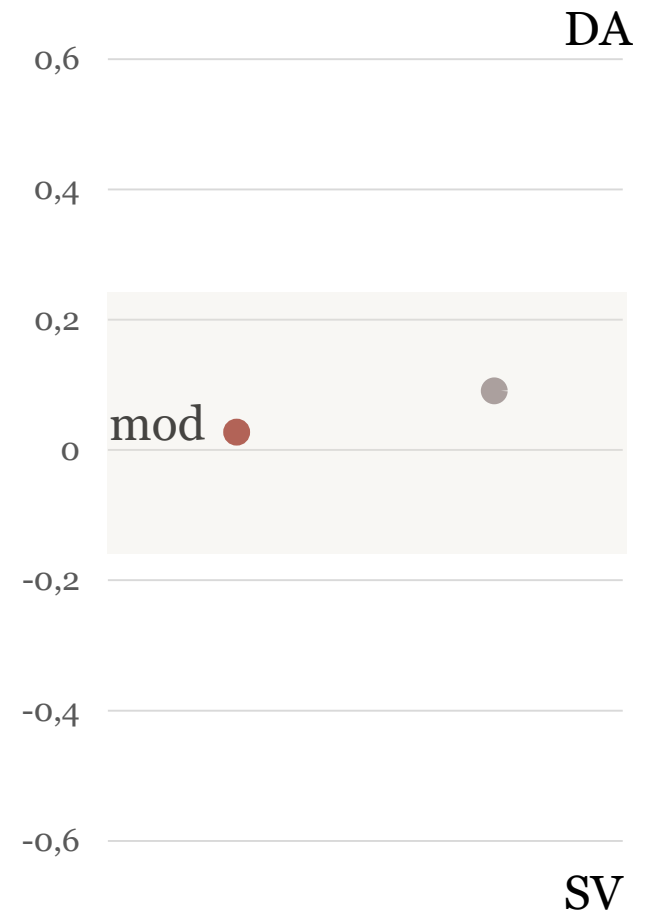
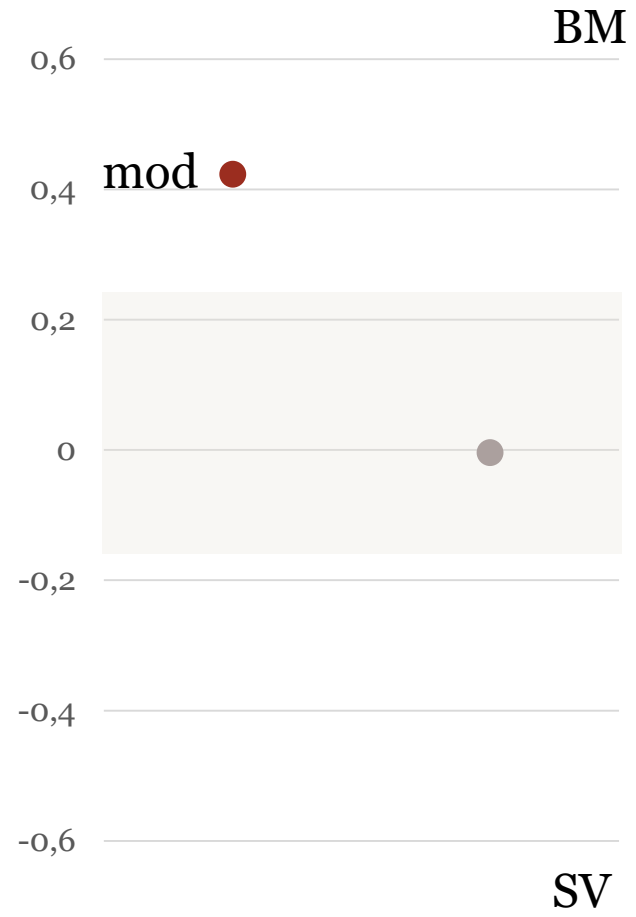
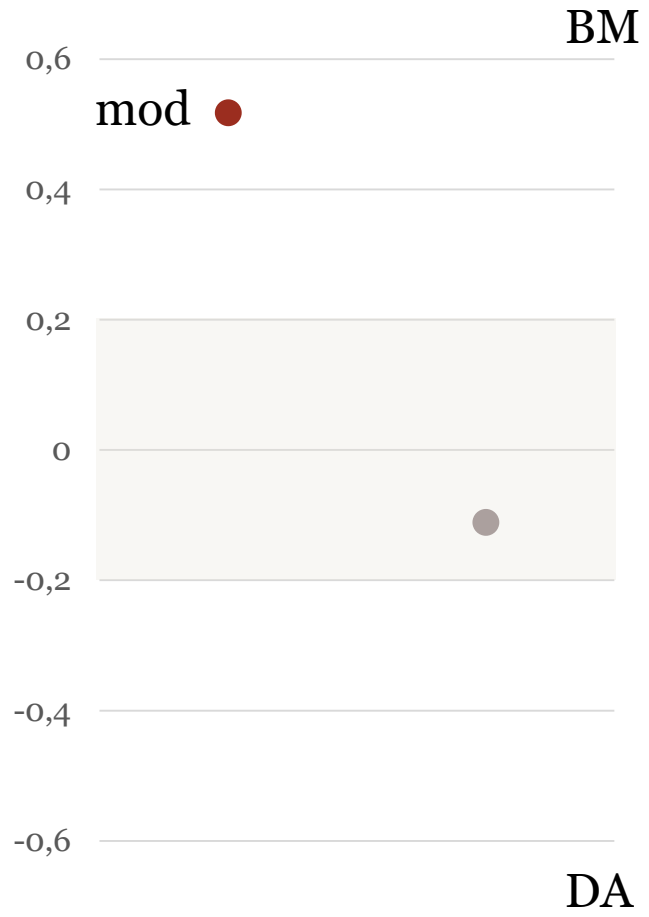




Han mannen







Prenominal modification

Swedish: *andra, gamle*, nationalities

- *Anyway, ska torka mina tårar och fortsätta tänka på **han gamle mannen**.*

Danish: primarily nouns/abbreviations

- *De var nu flinke allesammen – især **ham rulle-kebab manden**.*
- ***Ham IKEA manden** som var i lampeafdelingen mente at IKEA ville lave en master på et tidspunkt.*

Bokmål: much less semantic coherence

- *hvilken film likte Lewis Gilbert best å regissere? vel, ifølge ham selv er det "den med Roger Moore og **han gigantiske mannen** «*
- *jeg har aldri følt at **han activia mannen** var troverdig..*
- *å den lille ekstra strekningen fikk jeg tid til å tenke at nå tror vel både buss-sjåføren og **han glisende mannen** at jeg er en kjip tjei*

Conclusion

In sum

The differences between the neighborhoods, thusfar:

- DA: *ham manden's* form
- BM: *han mannen's* co-occurrence with prenominal modification
- BM: *han mannen's* co-occurrence with possessive pronouns?

- DA: *den mand's* form in Danish
- SV: *den mannen's* deictic reinforcement
- DA: *den mand's* association of RCs/ SV: *den mannen's* lack of association with RCs?

References

- Diessel, Holger. 2019. *The Grammar Network. How Linguistic Structure Is Shaped by Language Use*. Cambridge: Cambridge University Press.
- Greenwell, Brandon M. 2017. “Pdp: An R Package for Constructing Partial Dependence Plots.” *The R Journal* 9 (1): 421–36.
- Gries, Stefan Th. 2019. “On Classification Trees and Random Forests in Corpus Linguistics: Some Words of Caution and Suggestions for Improvement.” *Corpus Linguistics and Linguistic Theory*.
- Hothorn, Torsten, Kurt Hornik, and Achim Zeileis. 2006. “Unbiased Recursive Partitioning: A Conditional Inference Framework.” *Journal of Computational and Graphical Statistics* 15 (3): 651–74.
- Jakubíček, Miloš, Adam Kilgarriff, Vojtěch Kovář, Pavel Rychlý, and Vít Suchomel. 2013. “The TenTen Corpus Family.” *7th International Corpus Linguistics Conference*, 125–27.
- Johannessen, Janne Bondi. 2006. “Just Any Pronoun Anywhere? Pronouns and ‘New’ Demonstratives in Norwegian.” In *A Festschrift for Kjell Johan Sæbø*, edited by T. Solstad, A. Grønn, and D. Haug, 91–106. Oslo: University of Oslo.
- . 2008. “The Pronominal Psychological Demonstrative in Scandinavian: Its Syntax, Semantics and Pragmatics.” *Nordic Journal of Linguistics* 31 (02): 161–92.
- . 2020. “Psychologically Distal Demonstratives Are Not ‘Discourse New.’” In *Demonstratives in Discourse*, edited by Å. Næss, A. Margetts, and Y. Treis. Berlin: Language Science Press.
- Lie, Svein. 2010a. “Bakgrunndeiksis Og Demonstrativer.” In *Bo 65. Festschrift Til Bo Ralph*, edited by K. Jóhannesson, I. Larsson, E. Magnusson Petzell, S.-G. Malmgren, L. Rogström, and E. Sköldberg, 168–77. Göteborg: Meijerbergs institut.
- . 2010b. “Om Demonstrativer.” *Maal Og Minne* 2 (2010): 59–78.
- Strobl, Carolin, Anne-Laure Boulesteix, Thomas Kneib, Thomas Augustin, and Achim Zeileis. 2008. “Conditional Variable Importance for Random Forests.” *BMC Bioinformatics* 9: 307.
- Tagliamonte, Sali, and R. Harald Baayen. 2012. “Models, Forests, and Trees of York English: Was/Were Variation as a Case Study for Statistical Practice.” *Language Variation and Change* 24 (2): 135–78.