

Word order change in Germanic verbal clusters

Jelke Bloem, Arjen Versloot, Fred Weerman

The verbal cluster sandwich

■ Free order variation in Dutch

1. ik denk dat ik het **begrepen₂** **heb₁**
I think that I it **understood₂** **have₁**
2. ik denk dat ik het **heb₁** **begrepen₂**
I think that I it **have₁** **understood₂**

■ German, Frisian: Only **descending order**

■ English, Scandinavian: Only **ascending order** *

Verbal clusters

- Free order variation in Dutch
- German, Frisian: Only **descending order**
- English, Scandinavian: Only **ascending order** *
 1. I think that I **have**₁ **understood**₂ it

Why did they diverge?

English on one side, and German on the other

Modeling word order changes

- Can we model or simulate historical changes in verbal cluster word order?
- Start with proto-West-Germanic, end at the current state of the West-Germanic languages
- Language agents that produce and perceive verbal clusters

Language variation and change

How do we find factors involved in change?

- Language variation often caused by change

- Start by looking at the language with variation
 1. ik denk dat ik het **begrepen₂** **heb₁**
 2. ik denk dat ik het **heb₁** **begrepen₂**

- A language change in progress?

Correlates of variation: Meaning and function

- **Type of clause**
main clause / subordinate clause
- **Type of auxiliary**
copular-*zijn*/time/modal
- **Separable main verb**
... heeft afgewassen (has washed up)
- **Constituent after cluster**
... heeft gezien dat het gebeurde
- **Length of the middle field**
... dat [hij naar hun auto] is gelopen
- **Syntactic persistence**
...afgewassen heeft en ...weggelopen is
- **Main verb frequency**
... naar hun auto is gelopen
- **Multi-word units**
- **Pre-verbal constituent: Informativity and definiteness**

begrepen heb | heb begrepen

Storing preferences/frequencies

- Verbal cluster constructions in different contexts have different order preferences
 - Stored differently
 - Order preferences associated with context features?
- Store **red/green** order frequencies for each context feature (i.e. 'have'-cluster)

Model of a verbal cluster: Features

■ Type of auxiliary

- mod+inf ik denk dat ik het **zien wil**
- have+PP ik denk dat ik het **gezien heb**
- cop+PP ik denk dat hij **gezien is**

■ Type of clause: main clause/subordinate clause

1. This can not **be denied**.

■ 2 outcomes: **ascending** or **descending** order

■ How to model change?

Agent-based modeling of language

- Multiple language models that communicate
- Models a community of speakers (agents)
- Agents produce exemplar ‘sentences’ at each other
- Receiving agent learns from the sentence

- Also used in language evolution studies
 - Here the focus is not on communicative purpose, but on form and features

Agent-based modeling of language

- Simulates the fact that people do not perfectly copy a language from each other
- Learning bias -> change
 - i.e. deep structure bias
- Learning bias changes probability distributions

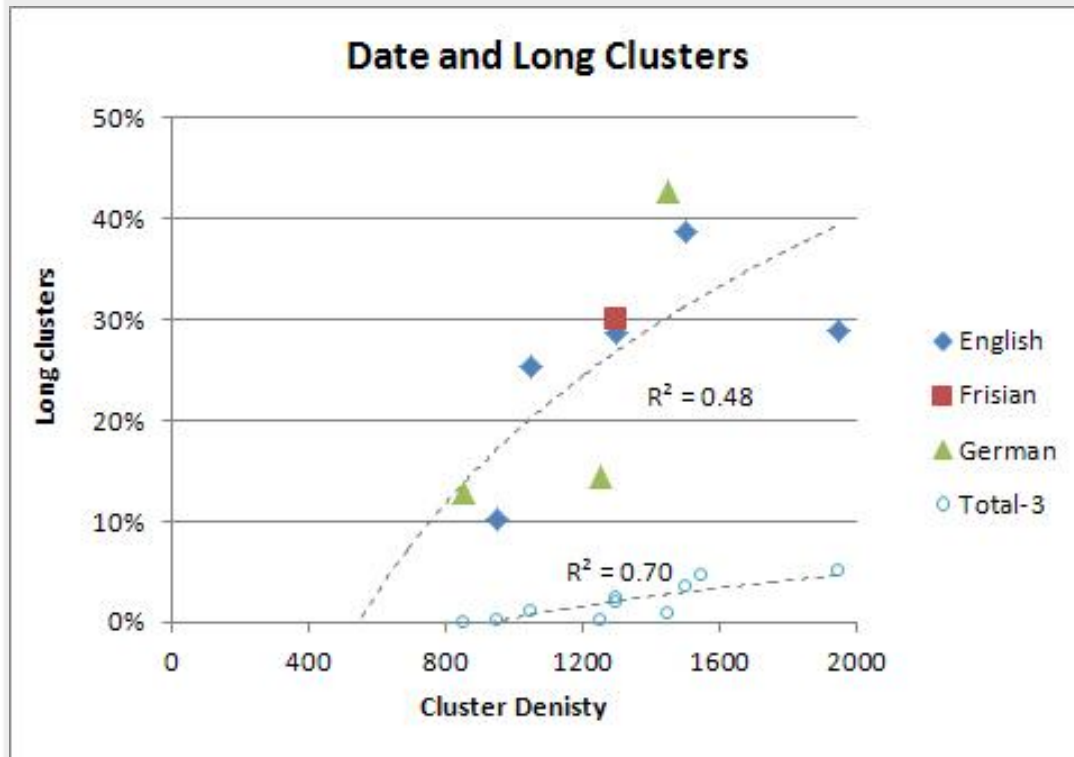
Modeling of agents (speakers)

- An example sentence looks like this:

	Modal + inf.	<i>to have</i> + part.	copula + part.
main clause	X		
subordinate clause			

- Create n agents with e exemplar sentences
- Next, agents transfer exemplars:
 - $p(\text{asc}|\text{mod-main}) = p(\text{asc}|\text{mod}) * p(\text{asc}|\text{main})$
 - Learning bias
- Starting situation based on proto-Germanic

Growth of multi-verb constructions in Germanic Languages



- The growth of 2-verb clusters in Germanic languages since ca. 500
- The growth of 3-verb clusters in Germanic languages since ca. 800.

Historical patterns underlying the model's starting position

- Constructions with *to have* growing from a very low level:

	Old	Modern
English: <i>have</i>	2%	31%
German: <i>haben</i>	1%	36%

- Emerged later than the first clusters, the modal+inf combination
- Growth phase in the model
- Increasing number of subordinate clauses

Outcome for 30 agents, 5000 interactions

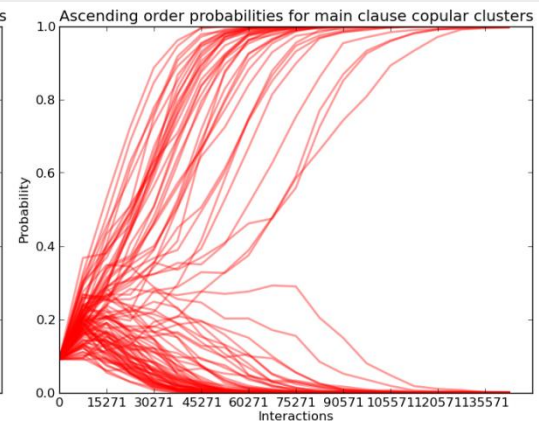
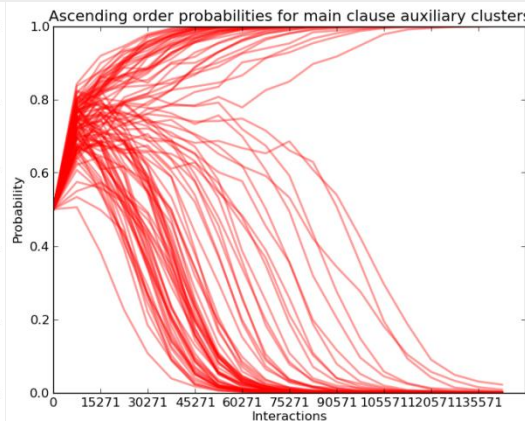
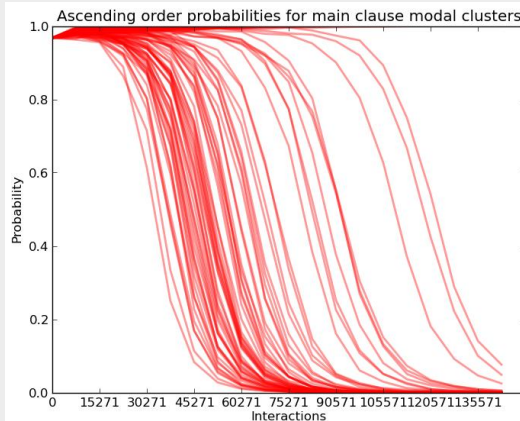
Equal increase of *to have*-constructions and subordinate clauses

Modal

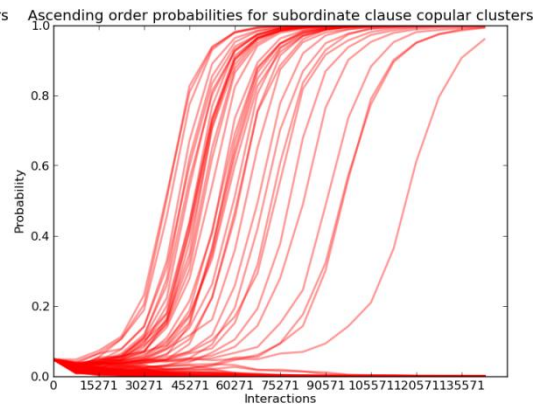
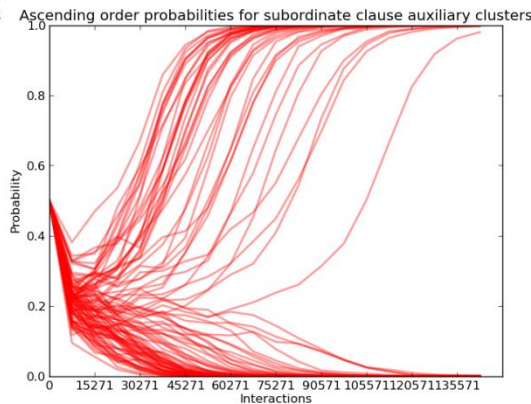
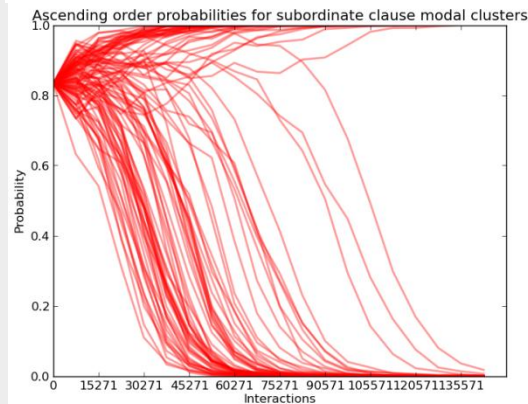
'to have'

Copular

MC



SUB



The model correctly predicts both dominant **ascending** (English) and **descending** (German)

begreifen heb

heb begreifen

Word order change in
Germanic verbal clusters

Outcome for 30 agents, 5000 interactions

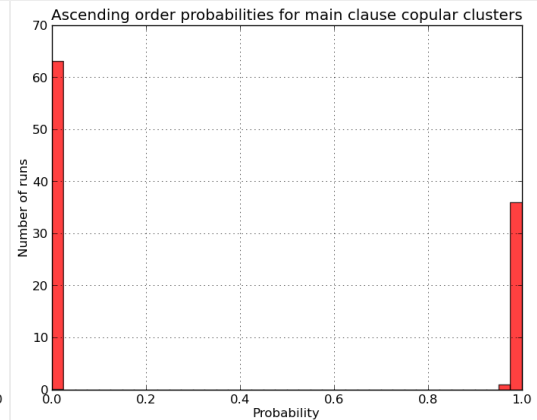
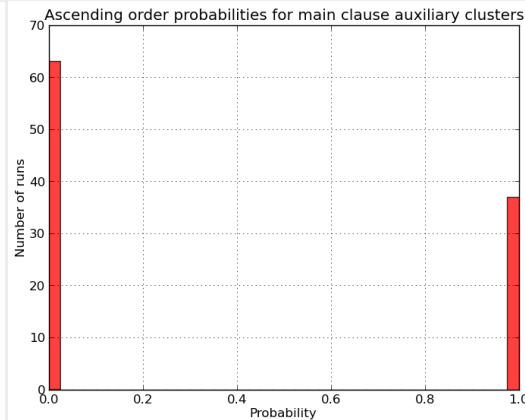
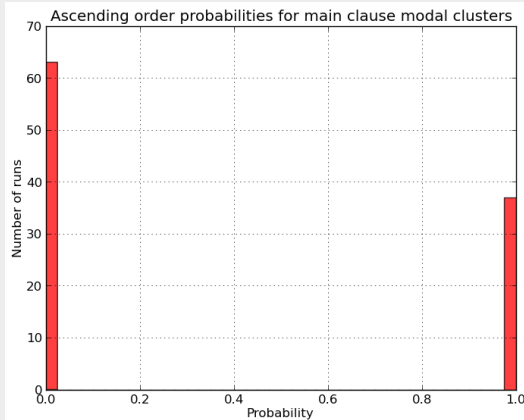
Equal increase of *to have*-constructions and subordinate clauses

Modal

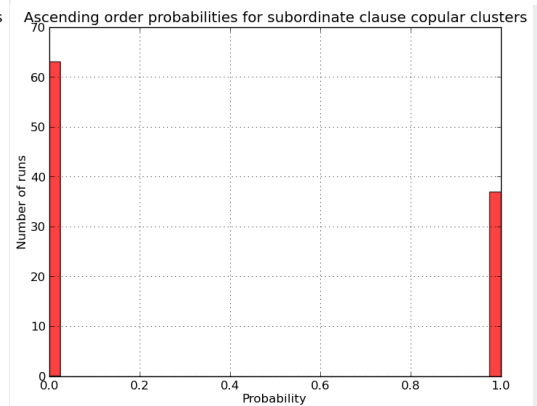
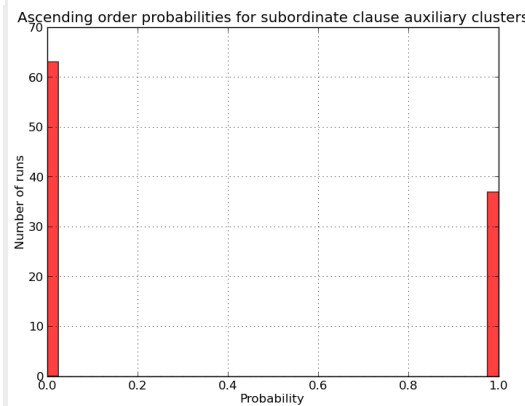
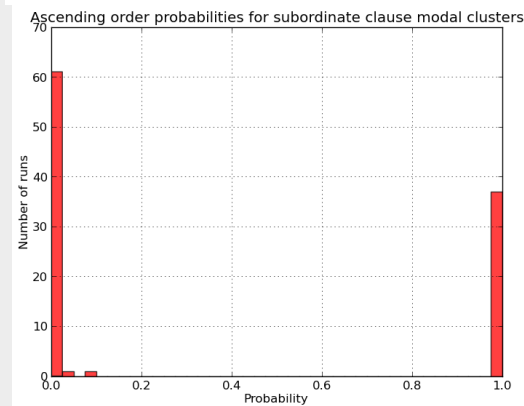
'to have'

Copular

MC



SUB



The model correctly predicts both dominant **ascending** (English) and **descending** (German)

begrepen heb | heb begrepen

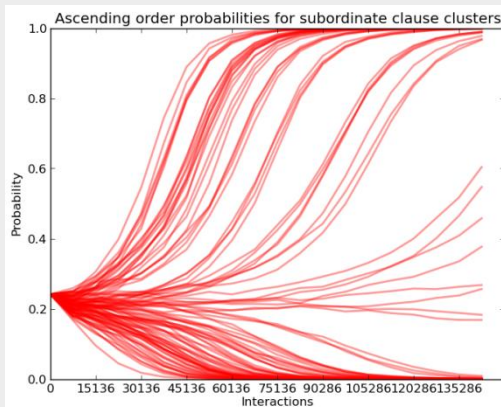
Word order change in
Germanic verbal clusters

Influence of the relative growth velocity of *to have*-constructions

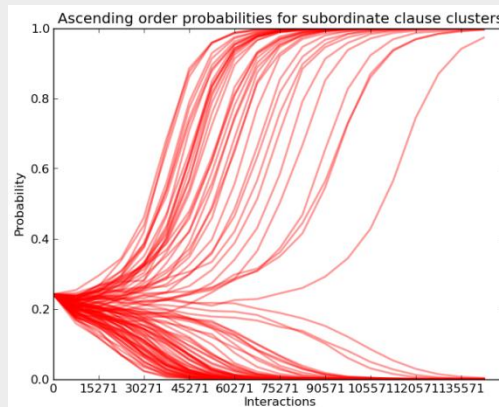
quick growth ('English')

moderate growth

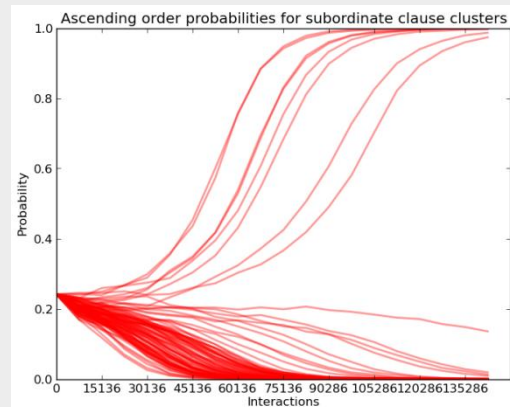
slow growth ('German')



56%/35%



63%/36%



92%/7%

Quicker 'have' growth increases the chances of an **ascending** word order

Results

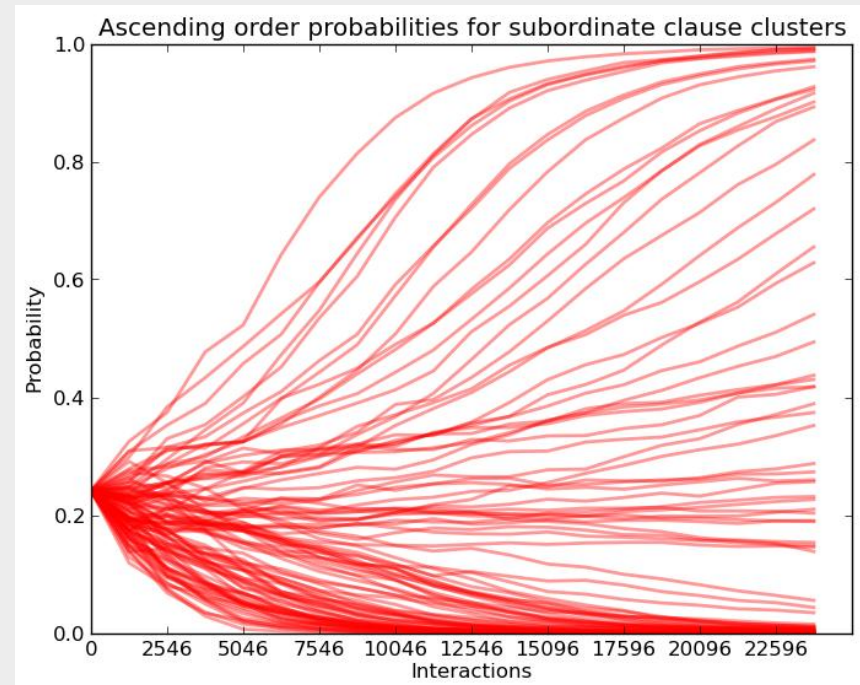
- Growth of 'have' supports **ascending** order
 - Prediction: more 'have' in English
 - Growth of subclauses supports **descending** order
 - Prediction: more sub clauses in German
- > The dominant word order may depend on different preference for specific constructions

Dutch variation: a change in progress?

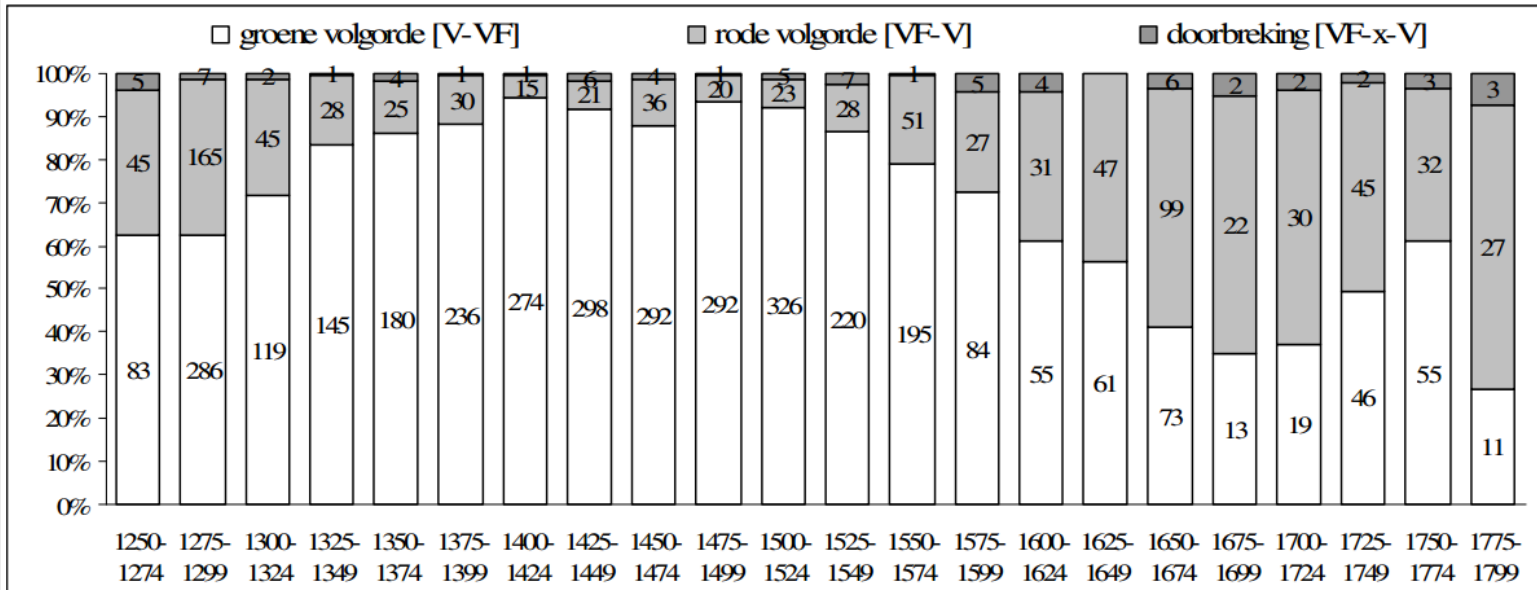
- Van Haeringen distribution: Dutch between German and English
 - Also seems to apply to verbal clusters...
- Dutch: German verb raising mechanism and **descending** order, English **ascending** order
- Model allows for this, but not forever

Dutch variation: a change in progress?

- Model may remain in unstable state for a while
 - Optionality
- Dutch is changing to 100% **ascending** order?



Dutch historical change



Verb order in official texts (n = 4327) (Coussé 2008)

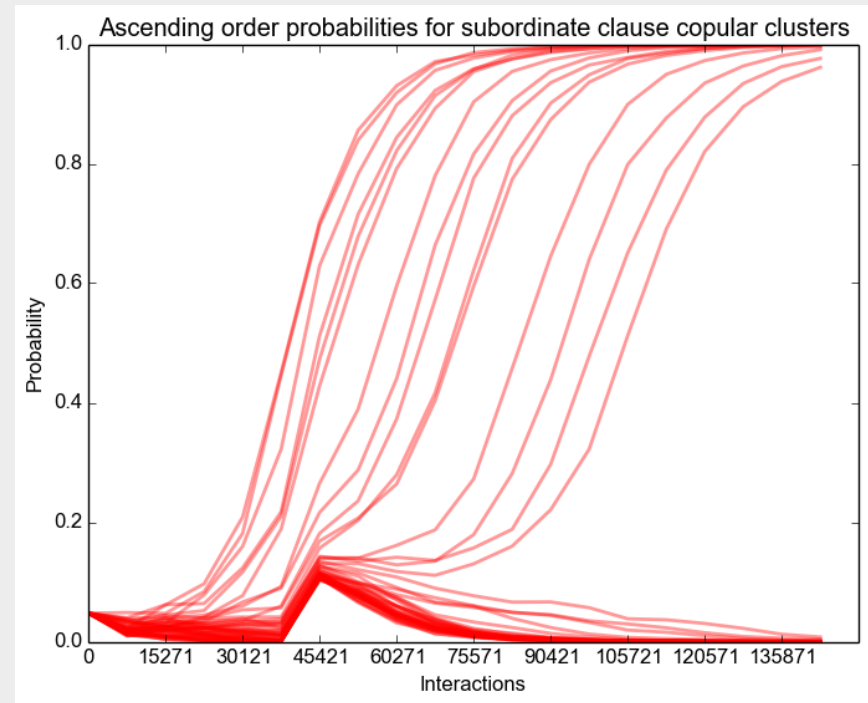
2 different processes of change?

Dutch: Another process of change?

- Dutch is changing to 100% **ascending** order?
 - **Ascending** order is acquired first (Meyer & Weerman, 2014)
 - **Ascending** order is catching on quickly in Frisian
- General preference for the **ascending** order?
(Not represented in our model)
- Frisian is changing due to language contact...
maybe Dutch too?

Dutch variation: language contact

- Crude simulation of language contact
 - Can cause change of default order
- Is this realistic?



Discussion

- Auxiliary type and clause type may be used as a diachronic explanation
- Increased use of subordinate clauses may have changed default order to **descending**
 - This can not **be denied**. (main clause)
 - ... that it not **denied** can **be**. (Contrasting)
- “Have” clusters support the opposite **ascending** order due to later grammaticalization
- Dutch requires a more complex model involving 2 processes of change

References

- J. Bloem, A.P. Versloot & F.P. Weerman (2014). [Applying automatically parsed corpora to the study of language variation](#). In *Proceedings of COLING 2014, the 25th International Conference on Computational Linguistics: Technical Papers* (pp. 1974-1984). Dublin, Ireland: Dublin City University and Association for Computational Linguistics.
- Coussé, E. (2008). *Motivaties voor volgordevariatie. Een diachrone studie van werkwoordvolgorde in het nederlands*. Universiteit Gent.
- Van Haeringen, C. B. (1956). *Nederlands tussen Duits en Engels*. Servire Den Haag.
- Versloot, A. P. (2008). Mechanisms of language change: vowel reduction in 15th century West Frisian.
- Wurmbrand, S. (2006). Verb clusters, verb raising, and restructuring. *The Blackwell companion to syntax*, 229–343.