

# Applying quantitative methods to dialect Dutch verb clusters

Jeroen van Craenenbroeck

KU Leuven/CRISSP

jeroen.vancraenenbroeck@kuleuven.be

## 1 Introduction

Verb cluster ordering is a well-known area of microparametric variation within Germanic (Barbiers & Bennis, 2010; Wurmbrand, 2005). For example, out of the six theoretically possible orderings in the three-verb cluster in (1), four are attested in Dutch dialects (Barbiers et al., 2008):

- (1) a. *Ik vind dat iedereen moet kunnen zwemmen.*  
I find that everyone must can swim  
'I think everyone should be able to swim.'
- b. *Ik vind dat iedereen moet zwemmen kunnen.*
- c. *Ik vind dat iedereen zwemmen moet kunnen.*
- d. *Ik vind dat iedereen zwemmen kunnen moet.*
- e. \**Ik vind dat iedereen kunnen zwemmen moet.*
- f. \**Ik vind dat iedereen kunnen moet zwemmen.*

In this talk I show that a quantitative analysis of these data can lead to new insights into the theory of verb clusters.

## 2 Methodology

I examine the raw data from 8 maps in the *Syntactic Atlas of the Dutch Dialects* (Barbiers et al., 2008): 4 containing two-verb clusters and 4 containing three-verb clusters, for a total of 28 possible cluster orderings. Based on Spruit (2008)'s version of the Hamming distance algorithm given in (2), I measure the differences in verb cluster ordering between 185 dialects of Dutch.

(2) **Hamming distance algorithm** (Spruit, 2008, p.36)

For each pair of dialects A and B, for each variant of all syntactic features, if it does occur in dialect A, but does not occur in dialect B or if it does not occur in dialect A, but does occur in dialect B, increment the distance between dialect A and B by 1.

The result is a (diagonally symmetric)  $185 \times 185$  matrix, a small portion of which is shown in figure 1, that lists, for each pair of dialects, the difference between them with respect to the 28 cluster orderings under investigation. This matrix was then analyzed using Multidimensional Scaling (implemented in R with `cmdscale`), which reduced the 185-dimensional variational space to a 2-dimensional one, in which each dialect received two coordinates representing its relative similarity to the other 184 dialects. The outcome of this procedure is shown in figure 2.

◇	A	B	C	D	E	F	G	H
1								
2	Boutersem	0	2	12	3	5	9	4
3	Wulvergem	2	0	14	5	5	7	4
4	Bakkeveen	12	14	0	11	13	11	12
5	Strijpen	3	5	11	0	6	12	1
6	Tongeren	5	5	13	6	0	8	5
7	Oldemarkt	9	7	11	12	8	0	11
8	Beveren	4	4	12	1	5	11	0
9	Sint-Gillis-Bij	1	3	11	2	4	10	3
10	Rotterdam	5	3	15	8	6	4	7
11	Roeselare	1	3	11	2	4	10	3
12	Waregem	2	2	12	3	3	9	2
13	Dilbeek	2	2	14	5	7	7	6

Figure 1: Cluster ordering differences between Dutch dialects

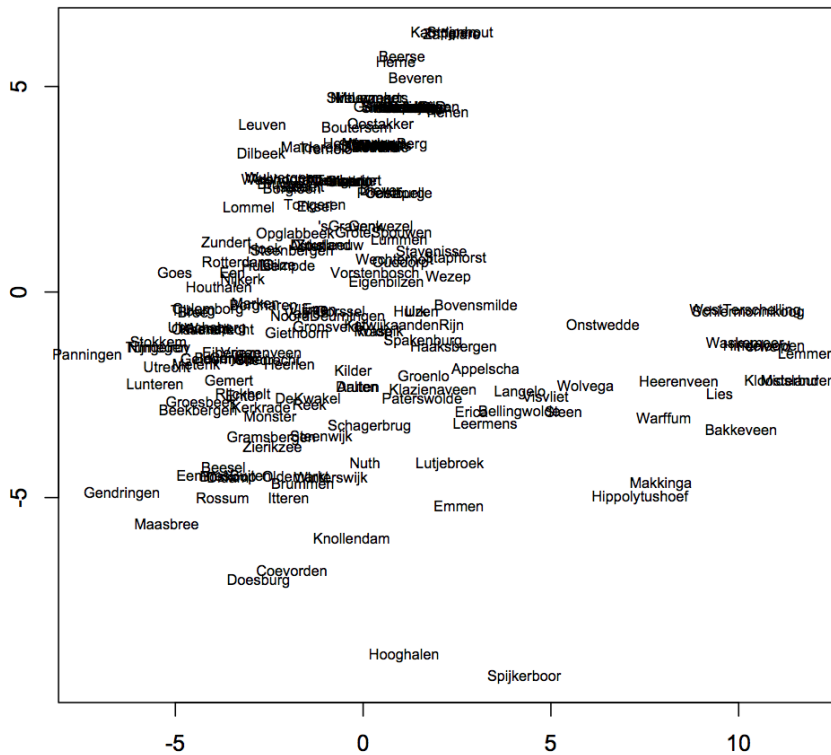


Figure 2: Two-dimensional representation of verb cluster ordering differences

### 3 Interpretation

Three clusters can be discerned in figure 2: those with an  $x$ -coordinate greater than 5, those with a  $y$ -coordinate greater than 2 and those in the range  $(-6,-5)$ - $(5,2)$ . Moreover, dialects with a  $y$ -coordinate below  $-5$  don't seem to pattern with any other dialects. Interestingly, if these groups are represented on a map of the Dutch language area, the result is surprisingly geographically homogeneous, see figure 3, suggesting that the patterns uncovered via this quantitative methodology correspond to an underlying linguistic reality. Preliminary investigations suggest that in the first pattern the main verb is placed at the left edge of the cluster, in pattern two a participle is placed to the left of its selecting verb, but an infinitive to the right, and in pattern three infinitives are placed to the right of their selecting verb, but participles either to the left or to the right.

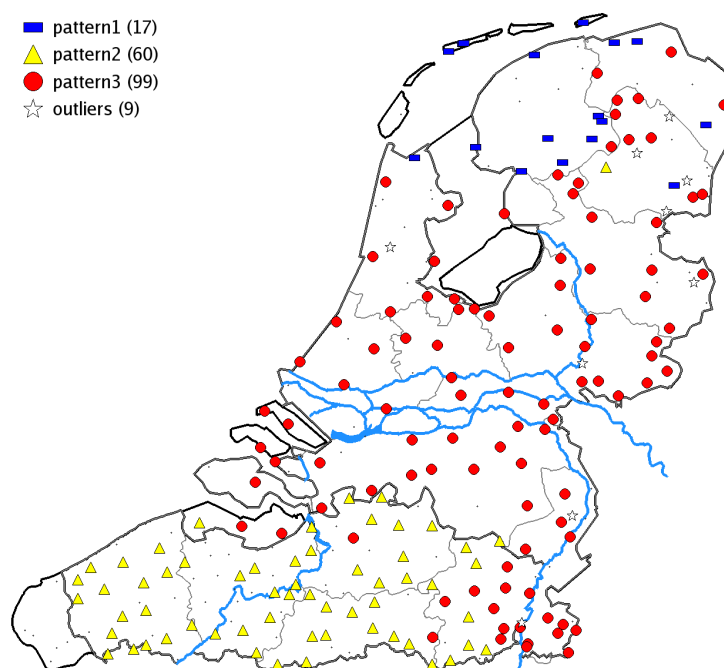


Figure 3: Verb cluster ordering patterns in Dutch dialects

## 4 Conclusion

This study shows that a purely quantitative methodology can shed new light on the theoretical analysis of verb cluster ordering in Germanic.

## References

- Barbiers, S., Auwera, J. v. d., Bennis, H., Boef, E., Vogelaer, G. D., & Ham, M. v. d. (2008). *Syntactische atlas van de Nederlandse dialecten. Deel II*. Amsterdam: Amsterdam University Press.
- Barbiers, S., & Bennis, H. (2010). De plaats van het werkwoord in zuid en noord. In J. D. Caluwe & J. V. Keymeulen (Eds.), *Voor Magda. Artikelen voor Magda Devos bij haar afscheid van de Universiteit Gent* (pp. 25–42). Gent: Academia.
- Spruit, M. R. (2008). *Quantitative perspectives on syntactic variation in Dutch dialects*. Unpublished doctoral dissertation, Universiteit van Amsterdam.
- Wurmbrand, S. (2005). Verb clusters, verb raising, and restructuring. In M. Everaert & H. v. Riemsdijk (Eds.), *The Blackwell Companion to Syntax* (Vol. V, pp. 227–341). Oxford: Blackwell.