Measuring correlations in dialect Dutch verb clusters

JEROEN VAN CRAENENBROECK
KU Leuven

Verb clusters in dialect Dutch form a well-known area of microvariation (cf. Wurmbrand 2005, Barbiers & Bennis 2010). For example, out of the six logically possible orders in three-verb clusters, five are attested, and their occurrence depends on the type of verbs, the hierarchical relations in the cluster, etc.

Missing from the literature, however, is a study of the correlations that exist between the various cluster orderings in the dialects of Dutch. In this study I examine the raw data from 8 maps in SAND II (Barbiers e.a. 2008): 4 containing two-verb clusters and 4 containing three-verb clusters, for a total of 28 possible cluster orders. Based on Spruit's (2008:36) version of the Hamming distance algorithm ("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.") I measure the differences in verb cluster ordering between the 267 SAND-dialects.

When zooming in on those clusters with a Hamming distance smaller than 4, three geographically distinguishable patterns of verb cluster ordering emerge (see below): (a) a first pattern found in the Dutch-speaking part of Belgium that places a participle to the left of its selecting verb, but an infinitive to the right, (b) a second pattern situated in the central and southern part of the Netherlands that places infinitives to the right of their selecting verb, but participles either to the left or to the right, and (c) a third pattern situated in the north of the Netherlands that places the main verb at the left edge of the cluster.