Explaining the speed of lexical change in historical Dutch

Karlien Franco[1] & Dirk Geeraerts[2]
[1]KU Leuven & FWO Flanders, karlien.franco@kuleuven.be [2]KU Leuven, dirk.geeraerts@kuleuven.be

Keywords: Dutch, diachronic change, lexical semantics, quantitative analysis, lexical replacement

Why are different words replaced by new synonyms at different rates? In some domains, new variants replace each other rapidly (e.g. *awesome* or *lit* to refer to something COOL), while other domains display more stability (e.g. EAR). Recent research has shown that the speed of lexical change is influenced by word-related features, like frequency, word class, length or age of acquisition (Bochkarev, Solovyev & Wichmann 2014; Monaghan 2014; Pagel, Atkinson & Meade 2007; Wichmann & Holman 2013). In this paper, we analyze whether characteristics of concepts play a role as well. Taking our lead from Franco et al. (2019) who showed that concept characteristics such as familiarity, vagueness and affect-sensitivity influence the amount of synchronic lexical variation in the base dialects of Dutch, we test whether these characteristics affect the speed of diachronic change in Dutch as well.

The data we use come from the *Middelnederlands Woordenboek* (Middle Dutch Dictionary: 1250-1550) and the *Woordenboek der Nederlandsche Taal* (Dictionary of the Dutch Language: 1500-1976), two large dictionaries of historical Dutch. We extract data from the digitized versions of these dictionaries with the DiaMaNT tool (Depuydt & de Does 2018), a semantic historical computational lexicon for Dutch, zooming in on 252 concepts from two semantic fields: body parts and clothing terms. In particular, for each body part or clothing concept we record all the variants that are available as (historical) synonyms to express the concept, as well as the times at which they were used (by relying on the citations available in the dictionary). For example, for the body part JAW, we record that it occurs with 5 variants between 1500 and 1550, including *kaak* (the current Standard Dutch lexeme), *kinnebak*, *pellorijn* and *kieuw*.

Next, we divide the dataset into 50-year periods. For each period, we calculate two types of information: (1) the number of variants in use at each time point, and (2) the proportion of variants in use at a given period that were also used during the previous period. Using this information, we can answer two research questions: (1) is the number of synonyms for the concept diachronically stable, or are there fluctuations (diachronic stability)?; (2) how quickly do variants disappear from the data, how quickly are they replaced with new synonyms (the rate of lexical replacement)?

Our hypotheses are that the factors that play a role in synchronic data, affect diachronic change as well, viz. familiarity, vagueness and affect-sensitivity. Moreover, we may find differences between the body part concepts and the clothing concepts as the former concepts have a higher degree of universality and may therefore only rarely be referred to with novel lexical items.

Preliminary results on the body part concepts indicate that there are some trends in the data that confirm the correlation between familiarity and diachronic stability on the one hand, and affect-sensitivity and diachronic stability on the other. For vagueness, the picture is less clear. Further data collection and analyses will take place in the coming months.

References

Bochkarev, Vladimir, Valery Solovyev & Søren Wichmann. 2014. Universals versus historical contingencies in lexical evolution. *Journal of The Royal Society Interface* 11(101). 20140841.

Depuydt, Katrien & Jesse de Does. 2018. The diachronic semantic lexicon of Dutch as Linked Open Data. In *Proceedings of the Eleventh International Conference on Language Resources and Evaluation (LREC 2018).* European Language Resources Association (ELRA), Paris, France, May 2018.

Franco, Karlien, Dirk Geeraerts, Dirk Speelman & Roeland Van Hout. 2019. Concept characteristics and variation in lexical diversity in two Dutch dialect areas. *Cognitive Linguistics* 30(1). 205–242.

Monaghan, Padraic. 2014. Age of acquisition predicts rate of lexical evolution. *Cognition* 133(3). 530–534.

Pagel, Mark, Quentin D. Atkinson & Andrew Meade. 2007. Frequency of word-use predicts rates of lexical evolution throughout Indo-European history. *Nature* 449(7163). 717–720.

Wichmann, Søren & Eric W. Holman. 2013. Languages with longer words have more lexical change. In Lars Borin & Anju Saxena (eds.), *Approaches to Measuring Linguistic Differences*, 249–284. Berlin: De Gruyter Mouton.