

Diasystematic construction grammar meets diachronic construction grammar meets constructicography: Towards a digital Afrikaans taboo constructicon

Monique Rabé & Gerhard B. van Huyssteen

North-West University, Potchefstroom, South Africa

{monique.rabe OR gerhard.vanhuysteen}@nwu.ac.za

Keywords: Afrikaans, constructicon, constructicography, constructional change, constructionalisation, etymology, language contact, multilingual, resource-scarce language, taboo language

One of the central ideas in construction grammar (CxG) is that our linguistic knowledge is structured as a mental inventory of constructions, i.e., a constructicon. Descriptive representations of such constructicons are best characterised as a blend between CxG and lexicography, and has come to be known as constructicography (cf. Lyngfelt, 2018:2). To our knowledge, no research has been done on how insights from diasystematic construction grammar (DCxG) can be captured in constructicographic resources; the same also applies diachronic construction grammar (DiaCxG). Therefore, the overarching question that we would like to raise in this contribution is:

How can information about multilingual language contact situations (DCxG information) that resulted in constructional change and constructionalisation (DiaCxG information) be captured in a digital representation of a constructicon?

For practical/applied purposes, we will focus on challenges arising in the development of a digital Afrikaans taboo constructicon. Although small-scale annotated taboo lexica already exist for a small number of languages, e.g., HurtLex (cf. Bassignana *et al.* 2018) and Taboo Wordnet (Choo and Bond, 2021), resources that capture information about taboo constructions larger than words, or information pertaining to their diachronic development (e.g., the etymology of words) do not exist. In addition to the need to address lacunae like these, our focus on Afrikaans foregrounds yet another challenge, namely data scarcity. Since Afrikaans is considered a resource-scarce language (Matthew, 2015:199), limited data exist for Afrikaans in general. This is even more so true about historical data pertaining to the development of taboo constructions.

To illustrate these challenges, we will use the Afrikaans word *fok* ('fuck') as a prototypical example of a taboo construction that developed in a multilingual society, specifically from Dutch *fokken*, and under influence of English *fuck*. We will pose some theoretical and methodological questions, and then attempt to present some preliminary, possible solutions for discussion during the workshop.

References

Bassignana, E., Basile, V. & Patti, V. 2018. HurtLex: a multilingual lexicon of words to hurt. In: Cabrio, E., Mazzei, A. & Tamburini, F., eds. *Conference proceedings*. 5th Italian Conference on Computational Linguistics (CLiC-it), Torino, Italy. pp. 1-6.

- Choo, Y.H.M. & Bond, F. 2021. Taboo Wordnet. In: Bosch, S., Fellbaum, C., Griesel, M., Rademaker, A. & Vossen. P., eds. *Conference proceedings*. 11th Global Wordnet Conference (GWC 2021), Potchefstroom, South Africa. Global Wordnet Association, University of South Africa (UNISA). pp.36-43
- Lyngfelt, B. 2018. Introduction: constructions and constructicography. In: Lyngfelt, B., Borin, L., Ohara, K. & Torrent, TT. *Constructicography: construction development across languages*. Amsterdam/Philadelphia: John Benjamins.
- Matthew, G. 2015. Using technology recycling to develop a named entity recogniser for Afrikaans. *Southern African Linguistics and Applied Language Studies*, 33(2):199-216.