

Dialect Data Integration: Paving the Way for New Research

Veronique De Tier, Katrien Depuydt, Jesse de Does

(Dutch Language Institute)

Keywords: Dialects, Dutch, Geo-mapping, Data visualization

The Southern Dutch dialect area consists of four dialect groups (Flemish, Brabantian, Limburgian and Zeeland dialects). For each of these dialect groups, a dictionary is available, accessible both online and in print. To obtain an integrated view on this Southern Dutch dialect area the Database of the Southern Dutch Dialects (DSDD) has been developed (Van Keymeulen e.a. (2019), De Tier e.a. (2021)). The aim of the platform is to integrate the data from all the available dictionaries and make them available for joint searches and further research. The results can also be visualised on a dialect map (De Tier (2023a)). The DSDD currently contains data from three onomasiological dictionaries. An extension of the database with data from semasiological dictionaries, starting with the Zeeland dialects, is currently under development (De Tier (2023b)).

Description of the data and the DSDD application

1) Onomasiological dictionaries:

The data consist of the dataset from the Flemish, Brabantian & Limburgian dictionaries, covering 30.000 concepts linking more than 500.000 dialect words. The DSDD enables one to search for individual dialect words, for all dialect words from a certain municipality/region, or for all words belonging to a certain concept. Filters are available to make the search query as accurate as possible (fig. 1). Based on the search results, interactive maps can be compiled automatically, visualising the data in its geographical setting. The results can be clustered in various ways, using frequency or etymology. The maps enable users to select different levels of detail to be visualised. A set of symbols and colours can be used according to the user's needs (fig. 2).

2) Semasiological dictionaries:

Two extensive alphabetical dialect dictionaries containing more than 4.000 dialect words not only include the dialect words and their meanings but also provide geographical information regarding the usage of Zeelandic dialect words (region indicators and/or place names). This feature makes the data, in principle, mappable, facilitating integration into the concept-based DSDD database. At this moment part of these dictionaries are already integrated in the DSDD.

At the conference we like to demonstrate the results of the project following its realisation including directions for future research:

- the structure of the database,
- the search facilities,
- the cartographic tool to visualise the dialect words and its printing facilities,
- the export facilities
- the methods used to integrate semasiological data
- the possibilities for dissemination of the data (export facilities)

- future research made possible after integrating the data of different dialect dictionaries,
- the importance of the project beyond Dutch dialectology.

The screenshot shows the DSDD search interface. At the top, there is a search bar with the placeholder 'Ik zoek...' and a magnifying glass icon. Below the search bar is a search input field containing 'prei'. Underneath the search input are two radio buttons: 'Dialectwoorden voor...' (selected) and 'Betekenissen van...'. Below these are two buttons: 'Zoeken' (Search) and 'Reset'. The main area is titled 'Filters' and contains several dropdown menus and checkboxes. The 'Woordenboek' section includes checkboxes for Brabants (WBD) (213.257), Limburgs (WLD) (210.782), and Vlaams (WVD) (160.713). The 'Land' section includes checkboxes for België (352.445), Frankrijk (13.719), and Nederland (258.063). The 'Thema' section has a dropdown menu set to 'Thema'. The 'Provincie' section has a dropdown menu set to 'Provincie'. The 'Plaats' section has a dropdown menu set to 'Plaats'.

Figure 1: search facilities of the DSDD

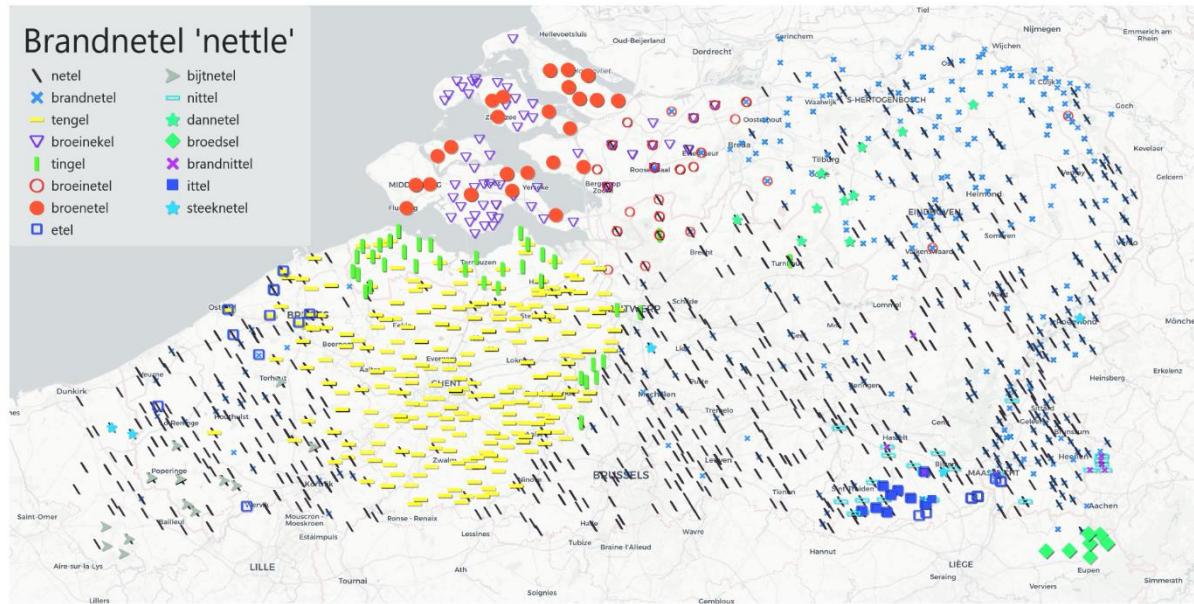


Figure 2: the words for nettle in the Southern Dutch dialects, including the Zeelandic data

References

Van Keymeulen, Jacques, V. De Tier, R. Vandenberghe en S. Chambers (2019), The Dictionary of the Southern Dutch Dialects (DSDD). Designing a Virtual Research Environment for Digital Lexicological Research, in Ed. Maria Pilar Perea. *Dialectologia - Special Issue 8*: 93–115.

De Tier, V., Chambers, S., Hellebaut, L., Van Keymeulen, J., Vandenberghe, R., Depuydt, K., De Does, J., Mertens, K. & Schoonheim, T. (2021), The Dictionary portal of the Southern Dutch Dialects, in Z.

Gavriilidou, L. Mitits, & S. Kiosses (Eds.), *EURALEX 19, Congress of the European Association for Lexicography, Abstracts*, 87-90.

De Tier, V. (2023a). 'Het Instituut voor de Nederlandse Taal leert nu ook dialect', in Sterkenburg, P. van, Bonth, R. de & Heylen, K. (red.), *In termen van taal*. Liber amicorum Frieda Steurs. [Schiedam], Scriptum, 317-323.

De Tier, V. (2023b). 'Zeeuwse dialectwoorden op het internet'. In: *Zeeuws Erfgoed*, jg. 22, 1, 26-27.