



# An ordinary linguist using Finnish corpora with Korp: Practical linguistics made easy

Mari Siirainen  
Department of Finnish, Finno-Ugrian and  
Scandinavian Studies

1. A linguist bumps into a word differing from the modern form of it in a Finnish text from 1651

Tuli cocht Cosee Corkee,  
Greitw' Pietari Rijkin Drotzet,  
Hai sield' tain **Neitzen** 'maiden' sorkeen.

'There came a high-ranking suitor, count Per Seneschal of the Realm,  
Who married this pretty maiden.'

Wijmeinen Muisto, [The Last Memory], part of a poem in memory of  
Kristina Katarina Stenbock, spouse of the the **count Per Brahe the  
Younger**, written by Ericus Justander



3. The linguist searches dialect corpora in the Korp

*neitte*

only 1 occurrence from south-western Finland

"**neitte** kulGi hevosen Gans ja hevone men suora päät  
lamme ja vei neitte mukanas" [Sauvo]

The modern form of the word being:

**neitsyt** 'virgin'

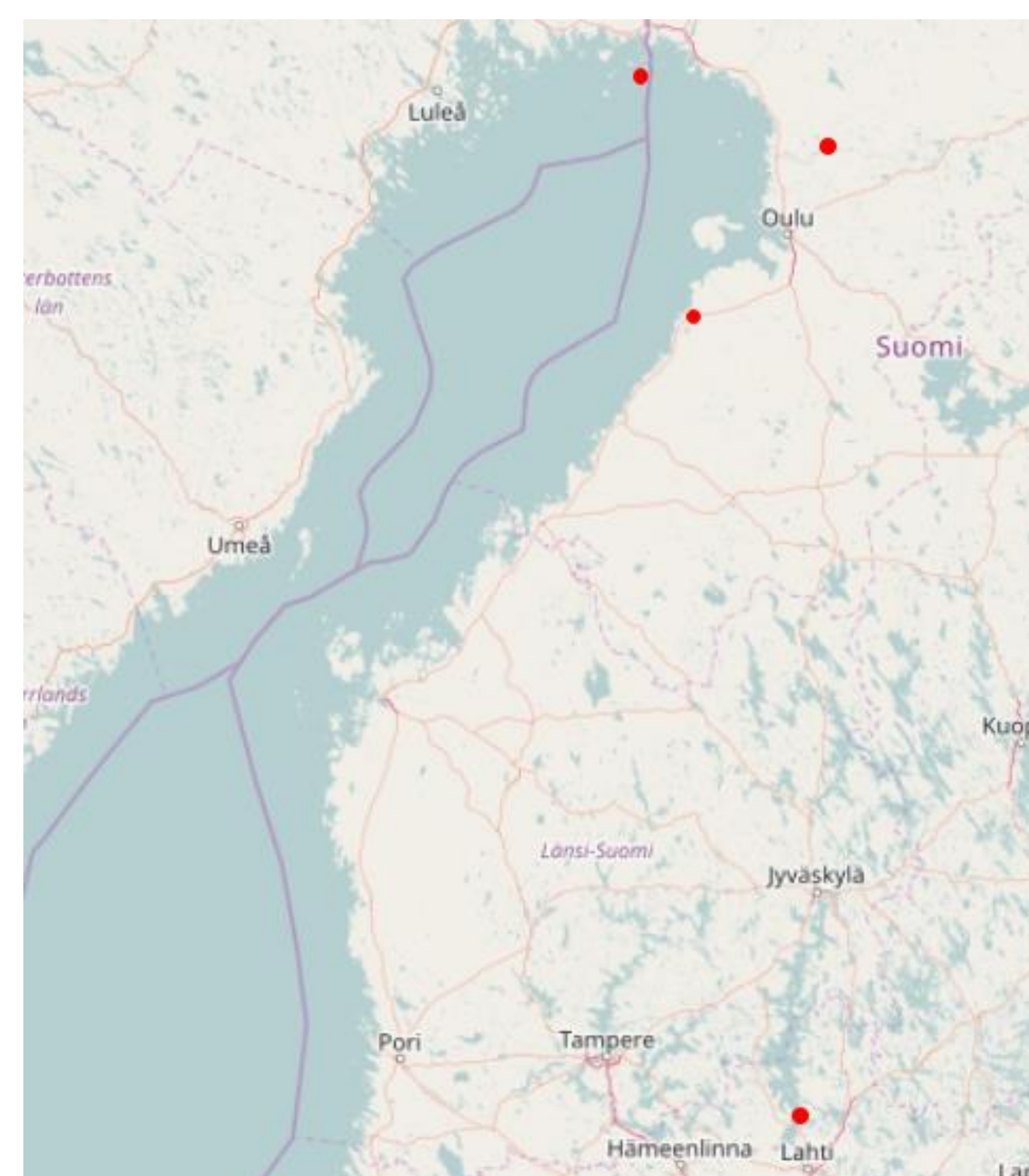
2. The linguist searches written corpora in the Korp

The word 'maiden, virgin' in Finnish texts  
before and after 1810

| word      | before 1810 | 1810 - 1870 |
|-----------|-------------|-------------|
| neitze    | 37          | 0           |
| neitzu    | 11          | 0           |
| neitzyt   | 15          | 0           |
| neitzö    | 6           | 0           |
| neidze    | 169         | 0           |
| neidzy    | 94          | 0           |
| neidzö    | 10          | 0           |
| neitse    | 0           | 28          |
| neitsy(t) | 0           | 198         |

*neitty*

5 occurrences from Northern Ostrobothnia



[ijoki]

Map directly from Korp

4. The research continues

Coming this far was very quick and easy thanks to Korp and all  
the corpora in it!