A standardized wordlist and new spellchecker of Frisian

Anne Dykstra, Pieter Duijff, Frits van der Kuip and Hindrik Sijens

1. Introduction
In 2015 the Frisian Language Web saw the light of day online (FLW, http://taalweb.frl/). It has been developed by the lexicography department of the Fryske Akademy and consists of a newly created standardized wordlist of Frisian, an online spellchecker, a machine translator and a dictionary portal. These four applications together and individually offer a unique language tool to help Frisian native speakers and learners to write proper Frisian. The FLW is the first step towards a Frisian language platform that makes use of modern web technology. The standardized wordlist of Frisian is an important part of FLW and will not only be a firm base for future dictionaries but also the backbone of a special spellchecker. In this article we will give a brief overview of the compilation process of the standardized wordlist and how it is applied in the spellchecker.

2. Standard Frisian
Frisian has two major dialects and a number of smaller ones. Traditionally, variant pronunciations from minor dialects are not considered standard and thus are not included in dictionaries (anymore)1. Nevertheless there is not a fully-fledged standard yet. Today’s standard is mainly based on the two major dialects: Clay Frisian and Wood Frisian. In practice the present dictionaries do not always make a clear and well-reasoned choice between the two. In daily practice it appears that the variant that in the dictionaries is given preference over Wood Frisian variants is ‘spacious, large, wide’, tommé (tomme ‘hat, cap’). Thus they have not been applied simultaneously:

(a) tradition
Though not fully-fledged, there is a certain standard Frisian language, which has been codified in Frisian dictionaries over the years. Though the dictionaries have most of the time not selected variants from peripheral dialects, they do not always make clear which variant from the two major dialects should be prioritized. Yet it has been common practice for a long time to consider the variants given first in the dictionaries as the preferred ones. That is why, for instance, Clay Frisian variants in -om or -omme (rom ‘spacious, large, wide’, tomme ‘thumb’) are given preference over Wood Frisian variants in -um of -üme (rûm, tûme).

(b) distancing
One element in the standardization of a language can be distancing, which implies moving away from another (dominant) language. In the case of Frisian, that other language is Dutch. An illustrative example is the preferred variant heel ‘whole’, though it is only used in a small part of Friesland. In by far the larger part of Frisians use heel, which is identical to the corresponding Dutch word and therefore disqualifies as a preferred variant.

Another example is the deletion of e in words like kiste ‘box’ and mûtse ‘hat, cap’.

3. Standard wordlist
The selection process for the wordlist has been guided by different criteria, though they have not been applied simultaneously:

1 Only the scholarly Wurdoek fan de Frysk-saal (http://gtb.inl.nl/) provides data from all dialects.
2 It must be said that there has also been resistance towards further standardization of Frisian, the main reason being fear of losing dialect variation.

Pieter Duijff is lexicographer at the department of linguistics of the Fryske Akademy. He was editor of a legal dictionary (2000), co-editor of the Wurdoek fan de Frysk-saal / Woordenboek der Friese taal (1984-2011), one of the editors-in-chief of the monolingual Frysk Hänwurdboek, and editor of word lists of Town Frisian dialects (1998) and modern Frisian (online, 2015). Since 2014 he is one of the editors of a new online Dutch-Frisian dictionary.

pduijff@fryske-akademy.nl

Anne Dykstra has recently retired from the Fryske Akademy, where he has taken part in various Frisian language projects, including the scholarly Wurdoek fan de Frysk-saal (Dictionary of the Frisian language). He was the editor of the Frysk-Ingelsk wurdboek (Frisian-English Dictionary), has worked on the Frisian Language Database, and was project manager of the Frisian Language Web. Currently, he is editor of the International Journal of Lexicography.
dykstraanne@gmail.com

Kernerman Dictionary News, June 2015
Frits van der Kuip has been a lexicographer at the Frysk Akademy since 1986. He was one of the editors of the 25-volume Wardboek fan de Frysk taal (Dictionary of the Frisian Language) and one of the editors-in-chief of the Frysk Hânwurdboek. In 2003 he completed his Ph.D. thesis on 17th century proverbs. More recently he was involved in the Frisian orthography and spellchecker project, and at present he is preparing an online Dutch-Frisian dictionary.

Hindrik Sijens is a lexicographer at the Frysk Akademy, and an editor of the Frysk Hânwurdboek (monolingual Frisian dictionary).