An Online Dictionary for Dialects of North Frisian

Tanno Hüttenrauch
Independent Researcher
Amsterdam, Netherlands
nurdfriisk@gmail.com

Michael Wehar
Swarthmore College
Swarthmore, PA USA
mwehar1@swarthmore.edu

Abstract

Language is an essential part of communication and culture. Documenting, digitizing, and preserving language is a meaningful pursuit. The first author of this work is a speaker of Söl'ring which is a dialect of the North Frisian language spoken on the island of Sylt in the North Frisia region of Germany. Söl'ring is estimated to have only hundreds of native speakers and very limited online language resources making it a prime candidate for language preservation initiatives. To help preserve Söl'ring and provide resources for Söl'ring speakers and learners, we built an online dictionary. Our dictionary, called friisk.org, provides translations for over 28,000 common German words to Söl'ring. In addition, our dictionary supports translations for Söl'ring to German, spell checking for Söl'ring, conjugations for common Söl'ring verbs, and an experimental transcriber from Söl'ring to IPA for pronunciations. Following the release of our online dictionary, we collaborated with neighboring communities to add limited support for additional North Frisian dialects including Fering, Halligen Frisian, Karrharder, Nordergoesharder, Öömrang, and Wiedingharder.

Keywords: Endangered Languages, North Frisian, Online Dictionary

1   Introduction

1.1   A Dialect of North Frisian

The North Frisian language is estimated to currently have 8,000 native speakers across many dialects including Fering, Halligen Frisian, Karrharder, Nordergoesharder, Öömrang, Söl'ring, Wiedingharder, and more (Salminen, 2007; Minderheiten in Schleswig-Holstein, 2021). The first author of this work is a speaker of Söl'ring which is a dialect that is spoken on the island of Sylt in the North Frisia region of Germany.

1.2   Our Goal

Our goal is to support North Frisian speakers and learners by creating and maintaining an online dictionary. Although existing organizations provide support for limited in-person programming and events to assist North Frisian speakers and learners, there is a lack of online resources and digital content for North Frisian dialects such as Söl'ring. In general, providing online resources and digital content for endangered languages is necessary not only for archival purposes, but also to support language speakers and learners.

2   An Online Dictionary (friisk.org)

2.1   Overview

We built our online dictionary as a web application called (friisk.org) to provide online resources and digital content for Söl'ring speakers and learners. Our web application was built with HTML, CSS, and JavaScript on the frontend and PHP, SQL, and Python on the backend. Our web application primarily functions as a dictionary and resource repository. The second author served as the primary developer for our web application.

2.2   Language Data and Guides

Our language data sets are primarily composed of word lists, translations, conjugations, and rewrite rules for IPA transcription. The first author served as the primary lexicographer for compiling, transcribing, and formatting our language data sets. Our language data sets were created using the first author's first-hand language knowledge and acquired information from established resources such as (Jörgensen et al, 1981; Kellner, 2006; Lorenzen, 1977; Möller, 1973). We acknowledge that, as best as we could, we received written permission from publishers and/or original authors in cases that we directly or indirectly incorporated existing language data to build our language data sets.

To accompany our language data, for each supported dialect of North Frisian, we developed a language guide that provides basic background and information on the dialect along with information on pronunciations, grammar, and external references.

2.3   Spell Checking

We use the Open Source Spell Checker which is a web-based clientside spell checker that was developed and released by the second author (Wehar, 2019). This spell checker is ideal for our purposes because it is simple, multilingual, and runs on the client rather than the server which allows our server to focus on serving content rather than running complex algorithms. This spell checker simply looks at heuristic character differences that are easy to check, but is optimized to check these differences across word lists containing up to 100,000 words.

2.4   Translations

We currently provide translations from German to Fering, Halligen Frisian, Karrharder, Nordergoesharder, Öömrang, Söl'ring, and Wiedingharder. Additionally, we provide translations from English to Söl'ring and from Söl'ring to German. In total, our data sets include over 75,000 translations.

For each language pair, we translate from common words or phrases to definitions or explanations. In some cases, we also include parts of speech and example phrases to further clarify the concept. Additional labels are provided in some cases to denote subdialect differences and differences between sources.

Language Pair	Number of Translations
english-solring	1,656
german-fering	3,326

german-halligen frisian	5,512
german-karrharder	3,321
german-nordergoesharder	3,433
german-oomrang	3,308
german-solring	28,395
german-wiedingharder	3,394
solring-german	24,709

Table 1: Total number of translations

2.5 Conjugations

We currently provide conjugations for over 3,000 Söl'ring verbs along with a small collection of conjugations for the other six dialects of North Frisian that we support.

Dialect	Number of Verbs Conjugated
fering	189
halligen frisian	125
karrharder	114
nordergoesharder	174
oomrang	120
solring	3,505
wiedingharder	146

Table 2: Total number of verbs conjugated

2.6 Pronunciations

We currently provide a beta system for automatic transcriptions from Söl'ring to the International Phonetic Alphabet (IPA). According to the first author, because Söl'ring was originally a spoken language rather than a written language, whenever a writing system was developed, spellings could be used to systematically infer pronunciation. In particular, sounds directly correspond to symbols (or pairs of sequential symbols) along with their context within a word. As a result, we were able to devise a list of over one hundred regular expression replacement rules (with varying priorities) to directly translate words into phonetic transcription. To verify the accuracy of the resulting transcriptions, a test set of word and transcription pairs were compiled by hand. It is important to note that there are regional differences in pronunciations that we hope to further investigate and support in the future.

3 Conclusion

3.1 Community Usage

We created an online dictionary to support North Frisian speakers and learners. Our online dictionary currently serves over 5,000 unique users per year and has been featured by multiple community organizations from the island of Sylt. Although our online dictionary primarily provides resources for the Söl'ring dialect of North Frisian, through interactions with online language communities, we have been able to support six additional dialects of North Frisian spoken in neighboring regions.

3.2 Future Work

We hope to continue to maintain our online dictionary for years to come to be able to continue offering online resources and digital content for speakers and learners of the dialects of North Frisian. We also hope that our model of providing language guides, spell checking, translations, conjugations, and pronunciations will be adopted by other language communities for language preservation initiatives. In addition, we plan to make some of our developed technologies available in open source repositories for free usage within other language preservation projects.

Bibliographical References

Jörgensen, V. T., Petersen, J., & Altstädt, C. (1981). Kleines Friesisches Wörterbuch der Nordergoesharder Mundart von Ockholm und Langenhorn : Huuchtjüsch, Freesch/Fräisch. Bräist: Nordfriisk Instituut.

Kellner, B. (2006). Sölring Uurterbok: Wörterbuch der sylterfriesischen Sprache. Söl'ring Foriining.

Lorenzen, J. (1977). Deutsch-Halligfriesisch : e. Wörterbuch : 6000 Vokabeln Halligfries. mit Texten aus d. 17. bis 20. Jh. = Tutsk-freesk : en üürdeböök. Bräist/Bredstedt: Nordfriisk Inst.

Minderheiten in Schleswig-Holstein - Friesen (2021). Retrieved on May 24, 2022 from https://www.schleswig-holstein.de/DE/Fachinhalte/M/minderheiten/minderheiten_friesen.html

Möller, B. (1973). Söl'ring Uurterbok: Wörterbuch der Sylter Mundart. Sändig.

Salminen, T. (2007). Europe and North Asia from: Encyclopedia of the world's endangered languages. Routledge.

Language Resource References

Wehar, M. (2019). Open Source Spell Checker. https://github.com/MichaelWehar/Open-Source-Spell-Checker