



LUND UNIVERSITY

Lexical proficiency as a measure of language vitality in North Sámi

Lloyd-Smith, Anika; Gyllstad, Henrik; Kupisch, Tanja

2021

Document Version:

Publisher's PDF, also known as Version of record

[Link to publication](#)

Citation for published version (APA):

Lloyd-Smith, A., Gyllstad, H., & Kupisch, T. (2021). *Lexical proficiency as a measure of language vitality in North Sámi*. Poster session presented at International Conference on Minority Languages XVIII, Bilbao, Spain.

Total number of authors:

3

General rights

Unless other specific re-use rights are stated the following general rights apply:

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal

Read more about Creative commons licenses: <https://creativecommons.org/licenses/>

Take down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

LUND UNIVERSITY

PO Box 117
221 00 Lund
+46 46-222 00 00

Lexical proficiency as a measure of language vitality in North Sámi

Anika Lloyd-Smith¹, Henrik Gyllstad² & Tanja Kupisch^{1,3}
 University of Konstanz¹, Lund University², UiT The Arctic University of Norway³



Background

PROFILE

The **Sámi languages** are spoken in Sweden, Norway, Finland and the Kola Peninsula (Russia). While all Sámi languages are endangered, **North Sámi** is the most widely-spoken, with an estimated 20,000 speakers, mostly in Sweden & Norway (Svonní, 2008).

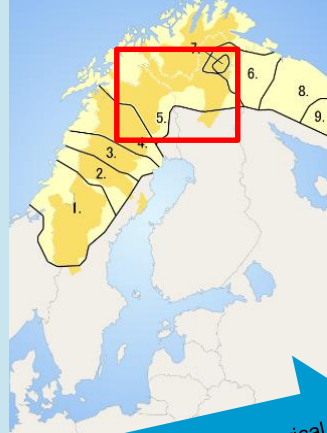
RESEARCH INTEREST

Since revitalizations efforts in the 1970s, Sweden and Norway have implemented **different policies** aimed at improving the **situation for Sámi**.

As part of an **interdisciplinary project** at the University of Konstanz, we investigate potential links between these policies and **language maintenance** practices, including:

- Language use
- **Lexical proficiency**
- Speech production
- Attitudes towards policies

Full data collection planned for **Spring 2021**



OBJECTIVE
 To develop an easy-to-use lexical measure of Sámi proficiency

The vocabulary measure

THE TEST FORMAT

- **Yes/No** vocabulary test following the DIALANG as a measure of general proficiency (e.g., Meara & Jones, 1987; Alderson, 2005; Lemhöfer & Broersma, 2012)
- Contains **50 real words** and **25 pseudowords** (verbs in infinitive form)
- Participants judge **whether word exists or not** (1 point/correct answer)

ITEM CREATION (FOR PILOT)

Real words (N = 100)

Verbs from Swedish and Norwegian DIALANG tests (from a range of frequencies) and translated into North Sámi. E.g., **šaddat** (to grow), **juohkit** (to divide)

Pseudowords (N = 50)

Constructed by native speaker of Sámi following phonotactic constraints of North Sámi. Often combining stem and ending of two verbs, e.g., **čalmmuhit** = **čatnat** (to connect) + **almmuhit** (to notify)

INCLUSION CRITERIA

- For inclusion in pre-test, the items needed to be
- Known to two native speakers and appear in SIKOR North Saami Corpus (real words)
 - Judged as plausible by two native Sámi speakers, with no possible meanings and not in SIKOR corpus (pseudowords)
 - Contain min. 5 and max. 12 letters (Table 1).
- Pre-test with **14 intermediate-native-like speakers of North Sámi** on SosciSurvey

Pre-test results

PRE-TEST (N = 14 participants)

Subsample was subjected to the following statistics (Table 1):

- **Point-biserial correlations** between response to each item and participants' total accuracy (cf. Amenta et al., 2020)
- **Item facility**, i.e., percentage of correct answers
- **Item discrimination**, i.e., ability of an item to discriminate between individuals (Alderson et al., 1995)
- **Correlations** with self-assessed proficiency (Figure 1).

The **75 best items** (50 real words + 25 non-words) were selected for the final version of the test.

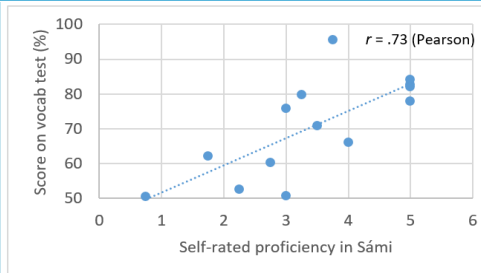


Figure 1. Relationship between vocabulary score and self-rated proficiency (reading/writing/listening/speaking) in Sámi

Table 1. Item statistics

	Length (letters)	Item facility (%)	Discrimination value	Point-biserial *correlations (r ²)
Pseudowords				
Min	6.00	14.29	0	.03
Mean	7.96	58.29	.62	.40
Max	12.00	100.00	1.00	.78
SD	1.86	18.03	.28	-
Realwords				
Min	5.00	50.00	0	.07
Mean	7.16	84.85	.66	.51
Max	12.00	100.00	1.00	.85
SD	1.56	12.01	.25	-

*correlation between item correctness and total scores/per participant

