Between standardization and variation:
Documenting and describing South Saami

In this talk I focus on ways to deal with variation attested in South Saami in my grammatical description for the language. While this variation is mainly based on geography and grade of grammaticalization, I illustrate how this language use contrasts with the standardized language of L2-speakers.

To set the scene, I first outline the conditions of describing South Saami and situate the language in its sociolinguistic context. South Saami belongs to the Uralic language family and is spoken in Norway and Sweden. It is a relatively small and endangered language with about 300–500 speakers. Possibly, there are only a couple of hundred (elderly) first language (L1-) speakers. These speakers are spread over a large area – a North-South dispersion of about 500 kilometres – and represent several different dialects. These dialects differ from each other in numerous features in phonology, morphology, syntax and lexicon. In addition, they differ also from the standardized language, which is in general more rigid, but also more conservative than the language of L1-speakers.

An initial challenge in the documentation project was to understand and navigate the complex mixture of different dialects/idiolects and standard language and become aware of influences from the latter. I decided to focus on spoken language of the older generation of L1-speakers. Focus on these speakers (arguably the last generation of L1-speakers in an unbroken tradition of South Saami) is also motivated by the variation in language use found in this generation. Examples are: allomorphs of the 1PL past tense subject suffix (*-mh, -bh, -o*); short forms of modal verbs used to mark future tense (*galkedh* 'shall' → *ga*; a particle-like use of the negative auxiliary; a less rigid word order (both SOV and SVO) – many of which had received little or no attention so far.

In order to verify that variation is not only idiosyncratic but systematic, I cross-checked the features in question with speakers from other dialects. Some features proved to be systematic and are used by most speakers of the older generation, across dialects. Other features appeared to be regional, but speakers from other dialects could sometimes comment that they recognized a particular feature from that dialect. The language of L2-speakers is mainly based on standardized South Saami. This (partly revitalized) language shows generally less variation and is for instance more rigid in word order than the language of L1-speakers and uses many fossilized inflectional suffixes productively.

Highlighting (systematic) variation in the language has therefore become an important part in the description – for descriptive and typological purposes, but also with revitalization of the language in mind: A corpus with examples for language in context might be a useful resource for second language speakers.

I believe South Saami constitutes an interesting case of a small minority language which is situated in a technological society while still being underdescribed. The language has an official orthography since 1978, but there is no comprehensive description of phonology that is agreed upon. The language undergoes revitalization and standardization and is used in social media. There are even language technology tools available (such as spell checks), while we still document and describe its grammar.