

## Undescribed and unstudied language resources of Leivu South Estonian

Janek Vaab

*Junior Research Fellow*

*University of Tartu*

janek.vaab@ut.ee

Leivu South Estonian, spoken near Alūksne and Gulbene in northeastern Latvia until 1988, represents the earliest branch of South Estonian (Kallio 2021). While the majority of existing descriptions of Leivu have been based on manuscripts, particularly the works of Valter Niilus (1935a, 1935b, 1937), audio recordings made between 1956 and 1962 have remained largely unexamined. To date, only 2 hours and 45 minutes of these recordings have been transcribed and published in a collection of Estonian dialect texts (Mets et al. 2014). Unfortunately, these transcriptions have been edited, with repetitions and self-repairs omitted, and they contain occasional inaccuracies. Those transcriptions account for only a small fraction of the over 12 hours and 8 minutes of recorded Leivu material housed in The Archive of Estonian Dialects and Finno-Ugric Languages (AEDFUL), nearly 10 hours of which remained untranscribed until 2023.

The aim of this presentation is to introduce the new on development morphologically annotated corpus of spoken Leivu South Estonian, based on these recordings. It outlines the need of the corpus, the transcription method used, and their importance in documenting this extinct language. Additionally, insights are provided into the speakers and recording conditions. The presentation addresses two main questions:

- (1) What linguistic features can be identified from the newly transcribed Leivu recordings?
- (2) How does Leivu compare to and align with other Southern Finnic languages, particularly Livonian and other South Estonian varieties?

Preliminary findings from the corpus reveal distinct linguistic features, including the phonetic variation of /j/, /h/, and /ʔ/, the use of the polar interrogative particle *vei(s)*, and debitive constructions (examples 1–3).

### (1) Leivu South Estonian

*nu ua `kiitmise `piehmest* (Mets et al. 2014: 95)

PP bean.PL **boil.DEB.PL** soft.TR

'The beans must be boiled until soft.'

(2) Livonian

*min'nən vɛ l' vīmāst sie rā' tā'giš* (Norvik, Tuisk 2024)

SG1.DAT be.PST **bring.DEB** DEM money back

'I had to bring this money back.'

(3) Seto South Estonian

*ma piüt rahha ka `kokko `korjama* (SetKo)

SG1 hold.SG2 money.PRT PP together collect.SUP.ILL

'I have to raise the money as well.'

The use of the debitive construction to mark deontic necessity sets Leivu apart from other South Estonian varieties and aligns it more closely with Livonian.

In conclusion, this research continues the work started in the 1950s offers a significant resource for the study of Leivu South Estonian, providing a dataset that can enhance our understanding of the language's structure. The creation of this corpus provides valuable data for future studies.

## References

**Kallio, Petri. 2021.** The Position of Leivu. *Eesti Ja Soome-Ugri Keeleteaduse Ajakiri. Journal of Estonian and Finno-Ugric Linguistics*, 12(2), 123–143.

<https://doi.org/10.12697/jeful.2021.12.2.05>

**Mets, Mari, Anu Haak, Triin Iva, Grethe Juhkason, Mervi Kalmus, Miina Norvik, Karl Pajusalu, Pire Teras, Tuuli Tuisk, Lembit Vaba. 2014.** *Lõunaeesti keelesaarte tekstid*. (Eesti murded IX). Tallinn: Eesti Keele Instituut, Tartu Ülikool.

**Niilus, Valter. 1935a.** Leivu murret: Häälikutelooline ülevaade. EMSUKA AES0179.

**Niilus, Valter. 1935b.** Leivu murret: Morfoloogiat (noomen ja verb). EMSUKA AES0178.

**Niilus, Valter. 1937.** *Valimik leivu murdetekste = Choix de textes dialectaux leivu*. Akadeemilise Emakeele Seltsi toimetised XXXI. Tartu.

**Norvik, Miina, Tuuli Tuisk. 2024.** Tartu Ülikooli liivi keele korpus. Tartu Ülikooli eesti ja üldkeeleteaduse instituut. [<https://doi.org/10.23673/re-473>].

**SetKo = Interdistsiplinaarne seto korpus SetKo.** <https://setko.ut.ee>