

Machine Translation for Livonian: Catering to 20 Speakers

Matīss Rikters^τ, Marili Tomingas^τ, Tuuli Tuisk^{τλκ}, Valts Ernštreits^λ, Mark Fishel^τ

^τ University of Tartu

{matiss.rikters, tuuli.tuisk, marili.tomingas, fishel}@ut.ee

^λ University of Latvia

valts.ernstreits@lu.lv

^κ University of Copenhagen

Abstract

Livonian is one of the most endangered languages in Europe with just a tiny handful of speakers and virtually no publicly available corpora. In this paper we tackle the task of developing neural machine translation (NMT) between Livonian and English, with a two-fold aim: on one hand, preserving the language and on the other – enabling access to Livonian folklore, lifestories and other textual intangible heritage as well as making it easier to create further parallel corpora. We rely on Livonian’s linguistic similarity to Estonian and Latvian and collect parallel and monolingual data for the four languages for translation experiments. We combine different low-resource NMT techniques like zero-shot translation, cross-lingual transfer and synthetic data creation to reach the highest possible translation quality as well as to find which base languages are empirically more helpful for transfer to Livonian. The resulting NMT systems and the collected monolingual and parallel data, including a manually translated and verified translation benchmark, are publicly released via the OPUS corpora collection and Huggingface model repository.

1 Introduction

Many state-of-the-art natural language processing tasks have reached admirable quality on languages with abundant linguistic resources (Vaswani et al., 2017; Conneau et al., 2018; Devlin et al., 2019). Furthermore, some neural language models and translation systems have been created for 100 and more languages (e.g. Conneau et al., 2020; Fan et al., 2021). However smaller, less or not at all spoken languages continue to struggle not only in terms of applicable computational approaches, but more critically - in terms of usable resources for training natural language processing (NLP) models or even just linguistic exploration.

In this paper we set the goal of developing machine translation between English and Livonian. Currently there are just over 20 fluent speakers of the language (Ernštreits, 2016). Although some digital linguistic resources exist for Livonian (including a dictionary with example sentences and a written monolingual corpus, Ernštreits, 2016), there is virtually no open parallel corpora between English and Livonian, with the single exception of 35 parallel sentences in the OPUS Tatoeba corpus (Tiedemann, 2020).

At the same time, cross-lingual transfer learning has recently helped improve the performance of several low-resource NLP tasks with the support of related languages (e.g. Conneau et al., 2018; Hu et al., 2020). This also includes zero-shot translation (Johnson et al., 2017), the ability of multilingual NMT systems to translate between seen languages that were not represented in the parallel training data as a pair. The case of Livonian is especially interesting in this regard, as there are two different sources of such support: on one hand, it is a Uralic language, closely related to Estonian and Finnish. On the other hand, Livonian has taken part in forming Latvian language and Livonian speakers have historically co-existed side-by-side with Latvian speakers. As a result of mutual influence these two languages also share a number of grammatical, lexical and orthographic similarities.

Our main contributions are two-fold. First, we collected the majority of digitally available translation examples including Livonian into a small parallel corpus (just over 10000 sentence pairs) of mostly Livonian-Latvian and Livonian-Estonian sentence translations with very few (1000) Livonian-English examples. In order to create a clean benchmark for evaluating translation quality we selected a portion (about 10%) of this corpus and had it manually translated into Latvian/Estonian/English so that each sentence would

Source	LIV-ENG	LIV-EST	LIV-LAT
Dictionary examples	–	10 690 / 44 854 / 44 499	10 690 / 44 854 / 44 975
Latvian constitution	686 / 11 198 / 15 499	719 / 11 454 / 10 314	719 / 11 454 / 11 002
JEFUL abstracts	–	187 / 2 878 / 2 846	176 / 2 723 / 3 434
Facebook posts	231 / 2 759 / 3 656	8 / 124 / 122	232 / 2 744 / 2 738
livones.net texts	169 / 2 741 / 3 660	92 / 1 969 / 1 867	333 / 4 449 / 4 433
Stalte ABC book	–	1 340 / 9 382 / 9 195	1 340 / 9 382 / 9 398
Trilium, poetry book	–	222 / 3 543 / 3 321	223 / 3 512 / 3 539
Eduard Vääri book	–	877 / 10 337 / 9 763	–
Total	1 086 / 16 698 / 22 815	14 135 / 84 541 / 81 927	13 713 / 79 118 / 79 519

Table 1: Total data size for the collected parallel LIV \leftrightarrow ENG/EST/LAT data. Each cell includes the sentence count, and word count for Livonian and the other language.

have all four manually verified translations.¹

The second half of our work focuses on neural machine translation (NMT, Vaswani et al., 2017), mainly targeting Livonian \leftrightarrow English. We explore several options of coping with the extremely low-resource settings and use Estonian and Latvian for cross-lingual transfer. Our experiments answer the following research questions:

1. Can we achieve machine translation for Livonian \leftrightarrow English at a usable level?
2. Which base language suits better for serving as base for cross-lingual transfer to Livonian, Estonian or Latvian?
3. Does zero-shot multilingual translation deliver better translation quality than pivot-translation through Estonian or Latvian?

Next we briefly describe the Livonian Language in Section 2, then introduce the collected parallel and monolingual data in Section 3. Section 4 provides the details of our NMT experiments and Section 5 concludes the paper.

2 The Livonian Language

Livonian (ISO 639-3: liv) is a Finnic language indigenous to Latvia and belonging to the Uralic language family. During the 12th century Livonian was spoken across great territories in Latvia around the Gulf of Riga. Over time, Livonian areas gradually became Latvian-speaking. In the 19th century, Livonian still had approximately 2500 speakers, by

¹Translation from Livonian was a too rare and expensive service, thus we resorted to translating from one of the other three languages and instead had Livonian speakers check the results for meaning correspondence afterwards.

the mid-20th century around 1500 speakers. Nowadays Livonian is listed in UNESCO’s Atlas of the World’s Languages in Danger as a critically endangered language (Moseley, 2014). According to the 2011 census, there are 250 Livonians in Latvia. Although there are just over 20 people who can speak the language, the Livonian community is active in preserving and developing the Livonian heritage (Ernštreits, 2016) and language plays a key role in this process (Ernštreits and Kļava, 2020).

The Livonian language developed in the contact area of Baltic and Finnic languages. Livonian and Latvian share a similar geographical location over a prolonged period of time, as a result of which they both contain traces of contact. Next to other loanwords, the Livonian loanword strata consists of words borrowed from Latvian (Suhonen, 1973; Winkler, 2014) and vice versa. The most obvious Latvian influence on Livonian grammar is found in the Livonian case system (Ernštreits and Kļava, 2014). Livonian has the prosodic characteristics typical of a Finnic language such as word-initial stress and the phonological opposition of short and long phoneme duration. It is the only Finnic language that differentiates lexical tones – the plain tone and the broken tone or *stød* – and therefore shares similar characteristics with Latvian as well as Danish (Tuisk, 2016).

3 Collected Data

The first step in developing (supervised) machine translation is collecting parallel data. While there was no pre-existing open parallel corpus with Livonian, we used all the possible sources of translations. This was limited to already digital resources, future work might include texts extracted by scanning older books and other materials.

	LV→EN	ET→EN	ETLV→EN	EN-ET-LV	Google	Neurotolge
ET		30.91	28.42	24.17	34.38	29.91
LV	25.18		25.26	20.77	31.54	25.92
LIV	2.20	3.22	2.66	13.29	-	-
Tuned						
LIV→EN	3.19	5.59	5.39	14.69	-	-
EN→LIV	-	-	-	8.59	-	-

Table 2: Results from machine translation experiments for translating into English. The source languages are listed in the first column and different models for translation are in each further column. We also compared ET/LV→EN translations of our evaluation set using Google Translate⁷ and Neurotolge⁸ online translation services.

The main sources of data included Livonian-Latvian as well as Livonian-Estonian translations. Thus we use these two languages as base for cross-lingual transfer and e.g. leave Finnish out, as there was no data for it.

The sources of data included:

- the Constitution of the Republic of Latvia, translated into 9 languages, including Livonian, Estonian and English,
- a database of dictionary entries, phrases and example sentences from the University of Latvia Livonian Institute’s website², with example sentences in Livonian, Estonian and Latvian
- the Livonian Institute’s Facebook page posts, partially parallel between our 4 languages
- books (Stalte, 2011; Kurs and et al., 2016; Ernštreit et al., 2020) with prefaces and content in Livonian-Estonian or Livonian-Latvian
- and abstracts from the Journal of Estonian and Finno-Ugric Linguistics’ (JEFUL) Special Issues on Livonian Studies (2014, 2016, 2018) in Livonian, Estonian and English.

Concerning sentence alignment, the dictionary examples consisted of already aligned Livonian sentences. We aligned the rest of the data manually with the help of language experts – first on paragraph level, then on sentence level. The resulting amount of sentences in the resulting dataset is shown in Table 1.

We separated balanced portions of development (503 sentences) and evaluation (749 sentences) splits from the full dataset. The splits are balanced in terms of the original source of the texts to resemble proportions from the remaining training data.

²www.livones.net/

We hired professional translators to create translations for any missing parts so that these splits would be parallel between all four languages. We further turned to experts of the Livonian language to make sure that the newly created translations truly convey the meaning of the original text as a quality control measure. The resulting benchmark and the whole corpus is published in the OPUS collection.³ We also share the final translation model⁴ after four iterations of backtranslation.

4 Machine Translation Experiments

Having just over 10,000 parallel examples constitutes extremely low-resource settings for neural machine translation. Added to this, the number of monolingual Livonian sentences (about 40,000) is also too small for approaches like unsupervised machine translation (Artetxe et al., 2018; Lample et al., 2018).

We implement the support of neighboring and related languages (Estonian and Latvian) via multilingual machine translation (Johnson et al., 2017). As a first step the model is pre-trained with the larger languages (Estonian, Latvian, English) and then used as base for following experiments.

We also perform iterative back-translation (Piniš et al., 2018) to make use of the large amounts of monolingual news data in EN/ET/LV, and our limited amount of monolingual data in LIV. We translate the 40k LIV sentences and different batches of 200k sentences from the other languages into all directions, filter the translations using simple heuristic filters (Riktors, 2018), and use a mix of all back-translated data with an equal amount of random clean parallel data (including all data involving Livonian) to fine-tune the base model.

³<https://opus.nlpl.eu/liv4ever.php>

⁴<https://huggingface.co/tartuNLP/liv4ever-mt>

	Base	Tuned	BT1	BT2	BT3	BT4
ET-EN	24.17	23.68	23.97	24.80	25.05	26.17
LV-EN	20.77	18.90	19.29	20.95	20.52	21.53
LIV-EN	13.29	14.69	16.19	17.41	18.15	19.01
EN-ET	17.00	16.87	18.58	19.37	18.95	19.48
LV-ET	18.38	19.55	19.72	19.93	20.68	22.38
LIV-ET	15.08	17.76	20.05	21.61	21.78	23.05
EN-LV	16.57	17.94	17.17	19.58	19.49	20.85
ET-LV	18.51	21.16	20.92	21.01	21.96	23.44
LIV-LV	15.05	17.55	21.25	22.99	23.68	25.24
EN-LIV	4.19	8.59	9.96	10.49	10.88	11.03
ET-LIV	4.01	13.00	14.43	15.24	16.09	16.49
LV-LIV	4.84	13.67	15.18	16.25	16.77	17.65

Table 3: Results in BLEU scores from the model at each training iteration translating in all translation directions.

4.1 Technical Setup

We used FairSeq (Ott et al., 2019) to train transformer architecture models with 6 encoder and decoder layers, 8 transformer attention heads per layer, word embeddings and hidden layers of size 512, dropout of 0.3, maximum sentence length of 128 symbols, and a batch size of 1024 words. All models were trained until they reached convergence (no improvement for 10 checkpoints) on development data. We used Sentencepiece (Kudo and Richardson, 2018) to create shared vocabularies of size 25,000, and SacreBLEU⁵ (Post, 2018) to generate BLEU scores (Papineni et al., 2002) for translations.

Base models were trained on LV→EN, ET→EN, ET+LV→EN data, and a multilingual model using the tagged approach (Johnson et al., 2017) for translating in all directions between EN/ET/LV languages. The base models were then used as initialization for tuning on Livonian-English parallel data.

For training the base models we used all available parallel data from Opus (Tiedemann and Nygaard, 2004). To facilitate further use of the base models for tuning on Livonian data, all Livonian sentences were used in addition to other data when creating the shared vocabularies. Finally, we used the highest-scoring tuned model to perform performed backtranslation on the monolingual LIV data to generate additional training data for training the final models.

4.2 Results

Table 2 shows the results of MT experiments. All BLEU scores are calculated for translations of our evaluation set. We compare the base single direction MT models to our multidirection model, as well as online translations from Google Translate⁶ and Neurotolge⁷ to evaluate performance from ET and LV into EN. While the multilingual model was noticeably weaker, the others hold comparable results to the online systems. However, when attempting to perform zero-shot translation from LIV into EN, ET→EN outperforms LV→EN (3.22 vs. 2.20), and the multilingual model achieved a very respectable BLEU score 13.29.

We then turned to tuning each of these models with LIV-EN data mixed 1:1 with a random equal amount of the original training data for each of the models. In the case of the multilingual model, we also added LV/ET-LIV data to the mix. This improved all scores by 1-3 BLEU points, but the multilingual model remained on top with 14.69 for LIV→EN. In order to perform back-translation models for both directions are required, so we scored the tuned multilingual model on the EN→LIV data as well, reaching 8.59 BLEU.

For comparison we also used the same tuned multilingual model to perform pivotal translation by first translating into ET or LV and then into the desired target language. In all four cases the pivot translation quality dropped when compared to direct translation by the same model, so we did not further pursue this line of experiments. An

⁵Case.mixed+numrefs.1+smooth.exp+tok.13a+version.1.5.1

⁶<https://translate.google.com> - accessed in Nov. 2021

⁷<https://neurotolge.ee> - accessed in Nov. 2021

	LIV-EN	EN-LIV
Facebook	19.28	13.55
Livones.net	19.67	15.91
Dictionary	7.73	10.60
Trilium	19.88	14.50
Stalte	13.88	9.47
JEFUL	8.02	5.10
Satversme	24.49	7.69

Table 4: Detailed experiment results in BLEU scores, split by the source of data from the last run of back-translation (BT4).

interesting observation, was that pivoting through ET achieved a higher BLEU score than LV when translating into EN (13.66 vs. 11.24), but slightly lower when translating into LIV (7.99 vs. 8.56).

Results for four rounds of BT iterations are compiled in Table 3. The model clearly improves not only in the main language pair of EN↔LIV, but in all other translation directions as well.

To answer the research questions, posed in the introduction, it seems that the resulting translation quality is still far from being usable. Comparisons between the base languages have shown slight preference towards Estonian over Latvian. Pivot-translation through Estonian or Latvian underperforms direct Livonian↔English translation trained in a zero-shot / few-shot manner.

4.3 Detailed Analysis

Table 4 shows BLEU scores of the separate parts of the evaluation corpus. Since most of the training data for EN-LIV comes from Satversme (Latvian Constitution), it is very clear why that part scores higher than others. The dictionary entries are overall far shorter in length than the other parts and often consist of few-word phrases, making them unfavorable to BLEU by definition.

The posts from Facebook and Livones.net are more general in their language and therefore more similar to data from the training set. However, the Trilium and Stalte books are written in a more literary language, making them slightly more challenging to translate. Finally, the very domain-specific part from JEFUL abstracts seems to be the most difficult to translate into English.

5 Conclusion

In this paper we presented a novel dataset for the highly endangered Livonian language, which can

be useful for machine translation, language modelling and many other natural language processing and computational linguistic research tasks.

In our experiments we show how far one can get in training modern machine translation models with very scarce data, and which languages are more suitable for transfer learning when working with Livonian data. While perhaps not being usable as-is in any kind of production scale, the achieved final BLEU scores of 19.01 for Livonian→English and 11.03 for English→Livonian show that some transfer of meaning can still be achieved with the currently available resources.

In the future we are planning to experiment with cross-lingual transfer from other languages, like the resource-rich Finnish as well as resource-poor Finno-Ugric languages like Võru and Sami (Tars et al., 2021). Given the limited amount of existing monolingual Livonian data, generating synthetic Livonian data with other means besides back-translation might be helpful: for example, forward-translation or using GPT-like language models.

Finally, work on the already collected Livonian monolingual and parallel data is ongoing at the Institute of the Livonian Language. Adding English translations to the lexical items and example sentences is an ongoing effort and will evaluate in practice, if the MT systems created as part of the current work can facilitate that. One of the key focuses is also manually verifying the data and making sure the existing corpus contains correct Livonian texts and their translations

Acknowledgements

This work has received funding from the “European Social Fund via IT Academy programme,” Estonian Ministry of Education and Research (grant SHVEE21397), the Estonian Research Council (grant GHVEE22112J), and the State Research Programme “Latvian Studies for the Development of a Latvian and European Society” project “Multifunctional dictionary of Livonian” (No. VPP-LETONIKA-2021/2-0002).

References

- Mikel Artetxe, Gorka Labaka, and Eneko Agirre. 2018. *Unsupervised statistical machine translation*. In *Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing*, pages 3632–3642, Brussels, Belgium. Association for Computational Linguistics.

- Alexis Conneau, Kartikay Khandelwal, Naman Goyal, Vishrav Chaudhary, Guillaume Wenzek, Francisco Guzmán, Edouard Grave, Myle Ott, Luke Zettlemoyer, and Veselin Stoyanov. 2020. [Unsupervised cross-lingual representation learning at scale](#). In *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*, pages 8440–8451, Online. Association for Computational Linguistics.
- Alexis Conneau, Ruty Rinott, Guillaume Lample, Adina Williams, Samuel Bowman, Holger Schwenk, and Veselin Stoyanov. 2018. [XNLI: Evaluating cross-lingual sentence representations](#). In *Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing*, pages 2475–2485, Brussels, Belgium. Association for Computational Linguistics.
- Jacob Devlin, Ming-Wei Chang, Kenton Lee, and Kristina Toutanova. 2019. [BERT: Pre-training of deep bidirectional transformers for language understanding](#). In *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long and Short Papers)*, pages 4171–4186, Minneapolis, Minnesota. Association for Computational Linguistics.
- Valts Ernštreits. 2016. Livonian in recent years. *Eesti ja soome-ugri keeleteaduse ajakiri. Journal of Estonian and Finno-Ugric Linguistics*, 7(1):257–274.
- Valts Ernštreits and Gunta Kļava. 2014. [Grammatical changes caused by contact between livonian and latvian](#). *Eesti ja soome-ugri keeleteaduse ajakiri. Journal of Estonian and Finno-Ugric Linguistics*, 5(1):77–90.
- Valts Ernštreits and Gunta Kļava. 2020. [Taking the livonians into the digital space](#). In *Proceedings of the 5th Conference Digital Humanities in the Nordic Countries*, pages 26–37, Riga, Latvia.
- Valt Ernštreit, Baiba Damberg, and Karl Pajusalu. 2020. *Trilium 2.0. Līvõ lūolkub. Liivi luulekogu. Lībiešu dzejas izlase*. The International Society of Livonian Friends.
- Angela Fan, Shruti Bhosale, Holger Schwenk, Zhiyi Ma, Ahmed El-Kishky, Siddharth Goyal, Mandeep Baines, Onur Celebi, Guillaume Wenzek, Vishrav Chaudhary, et al. 2021. Beyond english-centric multilingual machine translation. *Journal of Machine Learning Research*, 22(107):1–48.
- Junjie Hu, Sebastian Ruder, Aditya Siddhant, Graham Neubig, Orhan Firat, and Melvin Johnson. 2020. Xtreme: A massively multilingual multi-task benchmark for evaluating cross-lingual generalisation. In *International Conference on Machine Learning*, pages 4411–4421.
- Melvin Johnson, Mike Schuster, Quoc V. Le, Maxim Krikun, Yonghui Wu, Zhifeng Chen, Nikhil Thorat, Fernanda Viégas, Martin Wattenberg, Greg Corrado, Macduff Hughes, and Jeffrey Dean. 2017. [Google’s multilingual neural machine translation system: Enabling zero-shot translation](#). *Transactions of the Association for Computational Linguistics*, 5:339–351.
- Taku Kudo and John Richardson. 2018. [SentencePiece: A simple and language independent subword tokenizer and detokenizer for neural text processing](#). In *Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing: System Demonstrations*, pages 66–71, Brussels, Belgium. Association for Computational Linguistics.
- Ott Kurs and et al., editors. 2016. *Language and mind of the Livonian people. Eduard Vääri’s publications on Livonians and the Livonian language*, volume 74. Eesti Teaduste Akadeemia Emakeele Seltsi toimetised.
- Guillaume Lample, Myle Ott, Alexis Conneau, Ludovic Denoyer, and Marc’Aurelio Ranzato. 2018. [Phrase-based & neural unsupervised machine translation](#). In *Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing*, pages 5039–5049, Brussels, Belgium. Association for Computational Linguistics.
- Christopher Moseley. 2014. Livonian—the most endangered language in europe? *Eesti ja soome-ugri keeleteaduse ajakiri. Journal of Estonian and Finno-Ugric Linguistics*, 5(1):61–75.
- Myle Ott, Sergey Edunov, Alexei Baevski, Angela Fan, Sam Gross, Nathan Ng, David Grangier, and Michael Auli. 2019. [fairseq: A fast, extensible toolkit for sequence modeling](#). In *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics (Demonstrations)*, pages 48–53, Minneapolis, Minnesota. Association for Computational Linguistics.
- Kishore Papineni, Salim Roukos, Todd Ward, and Wei-Jing Zhu. 2002. Bleu: a method for automatic evaluation of machine translation. In *Proceedings of the 40th annual meeting of the Association for Computational Linguistics*, pages 311–318.
- Mārcis Pinnis, Matīss Rikters, and Rihards Krišlauks. 2018. [Tilde’s machine translation systems for WMT 2018](#). In *Proceedings of the Third Conference on Machine Translation: Shared Task Papers*, pages 473–481, Belgium, Brussels. Association for Computational Linguistics.
- Matt Post. 2018. [A call for clarity in reporting BLEU scores](#). In *Proceedings of the Third Conference on Machine Translation: Research Papers*, pages 186–191, Brussels, Belgium. Association for Computational Linguistics.
- Matīss Rikters. 2018. Impact of Corpora Quality on Neural Machine Translation. In *In Proceedings of the 8th Conference Human Language Technologies - The Baltic Perspective (Baltic HLT 2018)*, Tartu, Estonia.

- Kõrli Stalte. 2011. *Jelzi Sõnā: ābēd ja īrgandōks lugdōbrōntōz*. Jemākīel seļtš.
- Seppo Suhonen. 1973. Die jungen lettischen lehnwörter im livischen. *Suomalais-ugrilaisen seuran toimituksia*, 154.
- Maali Tars, Andre Tättar, and Mark Fišel. 2021. [Extremely low-resource machine translation for closely related languages](#). In *Proceedings of the 23rd Nordic Conference on Computational Linguistics (NoDaLiDa)*, pages 41–52, Reykjavik, Iceland (Online). Linköping University Electronic Press, Sweden.
- Jörg Tiedemann. 2020. [The tatoeba translation challenge – realistic data sets for low resource and multi-lingual MT](#). In *Proceedings of the Fifth Conference on Machine Translation*, pages 1174–1182, Online. Association for Computational Linguistics.
- Jörg Tiedemann and Lars Nygaard. 2004. [The OPUS corpus - parallel and free: <http://logos.uio.no/opus>](#). In *Proceedings of the Fourth International Conference on Language Resources and Evaluation (LREC'04)*, Lisbon, Portugal. European Language Resources Association (ELRA).
- Tuuli Tuisk. 2016. [Main features of the livonian sound system and pronunciation](#). *Eesti ja soome-ugri keeleteaduse ajakiri. Journal of Estonian and Finno-Ugric Linguistics*, 7(1):121–143.
- Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N Gomez, Łukasz Kaiser, and Illia Polosukhin. 2017. [Attention is all you need](#). In *Advances in Neural Information Processing Systems*, volume 30.
- Eberhard Winkler. 2014. [Loanword strata in livonian](#). *Eesti ja soome-ugri keeleteaduse ajakiri. Journal of Estonian and Finno-Ugric Linguistics*, 5(1):215–227.