

*ILSE LEHISTE* (Columbus)

## THE FINNO-UGRIC PROSODY PROJECT

### 1. Introduction

The purpose of this paper is to give a brief overview of a project a group of us have been engaged in for the past six or seven years — the Finno-Ugric Prosody Project.

The prosodic structure of three of the Finno-Ugric languages — Finnish, Estonian, and Hungarian — has been studied for some time and is relatively well known. The aim of the Finno-Ugric Prosody Project is to analyze the prosodic structure of less-known Finno-Ugric languages, using acoustic phonetic techniques that have not been applied to the study of these languages before. The core group of researchers is affiliated with Tartu University in Estonia and directed by Professor Karl Pajusalu, but scholars associated with other institutions are also involved.

We have been working more or less intensively for about seven years. Up to now we have completed our studies of the prosody of Erzya and Meadow Mari, and are in the process of analyzing Livonian. The selection of languages has been influenced by availability of speakers. The methodology involves recording of a corpus of systematically designed sentences by a set of speakers, and spectrographic analysis of these materials. The aim is to produce similar sets of data to facilitate comparison between languages.

According to the traditional view, a typical Finno-Ugric language is a language with word-initial stress and contrastive quantity. In Erzya and Mari, however, we have found neither contrastive quantity nor fixed initial stress. The phonetic feature most reliably associated with stress turned out to be lack of vowel reduction. There are, however, significant differences between these two prosodic systems; our materials make it possible to pinpoint these differences in a quantitative way. Livonian does have fixed initial stress, but it differs from the Finno-Ugric stereotype by having additional prosodic features often referred to as tones. We are planning to compare the Livonian prosodic system with the prosody of the neighboring Latvian dialect to explore possible similarities and differences.

Up to now, four publications<sup>1</sup> have resulted from the project: presentations at the 9<sup>th</sup> International Congress of Finno-Ugricists held in Tartu in 2001 on Erzya and Mari (published in the Proceedings), and books dealing with Erzya prosody (Helsinki 2003) and Meadow Mari prosody (Tallinn 2005). A summary of the Mari findings was also presented at the 10<sup>th</sup> International Congress of Finno-Ugricists, held in Yoshkar-Ola (the capital of Mari El) in 2005. Study of the prosody of Livonian is currently in progress. A summary of our Erzya and Mari results is offered below.

Our expectations about Finno-Ugric prosody are usually based on the prosodic systems of Finnish, Estonian and Hungarian, which are surprisingly similar. They all have phonological distinctions between short and long phonemes and no tonal contrasts that are independent of other prosodic features. They are all characterized by fixed word stress that does not, as a rule, distinguish between lexical and grammatical meaning. In native vocabulary, primary word-level stress is assigned to the first syllable, and secondary stresses to successive odd-numbered syllables. Tone is not independently contrastive, although it plays a part in distinguishing between long and overlong quantities in Estonian. All these languages have dynamic word stress, characterized by greater intensity and loudness of the stressed syllable.

Now Erzya has been claimed to have free stress, and some linguists have argued that there is no prominent word stress in Erzya at all. And according to a widely accepted point of view, Erzya has neither contrastive quantity nor tone. To be sure, the diachronic aspects of Erzya prosody have been extensively discussed and debated, but the synchronic descriptions are almost completely subjective, being based on auditory impressions. This is the gap our project is attempting to fill.

## **2.1. Erzya prosody: Materials and method**

During our study, we analyzed a set of 104 test words, placed in a frame sentence of the following type: "I said ..., not ... .", produced by eight speakers, four males and four females. The sentences were produced twice, the second time in reverse order, so that there were 104 occurrences of the test words in the middle of the sentence, and 104 occurrences in sentence-final position — for a total of 1632 productions. The test words were selected by Tiit-Rein Viitso, Professor of Finno-Ugric linguistics at Tartu University.

---

<sup>1</sup> The four current publications resulting from the Finno-Ugric project are the following:

I. Lehiste, E. L. Asu, E. Meister, K. Pajusalu, M. Parve, P. Teras, T.-R. Viitso, Acoustic Analysis of Erzya Prosody. — CIFU IX. Pars V, pp. 248—255.

I. Lehiste, E. Meister, K. Pajusalu, M. Parve, P. Teras, T.-R. Viitso, Acoustic Analysis of Meadow Mari Prosody. — CIFU IX. Pars V, pp. 256—263.

I. Lehiste, N. Aasmäe, E. Meister, K. Pajusalu, P. Teras, T.-R. Viitso, Erzya Prosody, Helsinki 2003 (MSFOu 245).

I. Lehiste, P. Teras, T. Help, P. Lippus, E. Meister, K. Pajusalu, T.-R. Viitso, Meadow Mari Prosody, Tallinn 2005 (Linguistica Uralica. Supplementary Series / Volume 2).

Seven of the speakers were born in Mordovia, one in Chuvashia; all had received their secondary and higher education in Mordovia, and were either students or faculty members at the State University of Mordovia. Their birth years ranged from 1940 (for a professor at Mordovia State University) to 1980 (for a student). Thus at the time of the recording (in the year 2000), the oldest subject was 60 years old and the youngest — 20.

The recordings were made by Einar Meister at the Laboratory of Phonetics of the Institute of Cybernetics of the Tallinn Technical University and at the Department of Estonian and Finno-Ugric Linguistics of the University of Tartu. The tapes were analyzed at Tartu by a team of graduate and doctoral students, using the Kay Elemetrics Computerized Speech Laboratory. Detailed measurements were made of various aspects of the signals; the results were treated statistically, and the results are published in the volume titled "Erzya Prosody" (2003; see Footnote 1).

## **2.2. Erzya prosody: Results**

We were looking for phonetic realizations of possibly contrastive prosodic features. It goes without saying that every sound in spoken language has a certain duration; unless it is whispered, it has a pitch, and in order to be produced and perceived at all, it has to have a certain amount of prominence. In our consideration of prosody, we are trying to identify and eliminate all differences between sounds that are externally conditioned, since such differences are not independently contrastive.

In the case of Erzya, we have identified a certain amount of conditioned variation. As regards quantity, for example, syllable nuclei were shorter in a closed syllable than in an open syllable. This suggests a certain amount of syllabic isochrony in Erzya. This durational difference is conditioned by syllable structure and does not constitute evidence for the existence of a prosodic opposition of quantity. The form of Erzya spoken by our subjects contains simple syllable nuclei (vowels) and complex syllable nuclei (diphthongs or glides); the complex nuclei are statistically reliably longer than simple syllable nuclei, but this difference is conditioned by the segmental structure of the syllable nuclei and therefore does not constitute evidence for the existence of a prosodic quantity opposition in Erzya.

As regards consonant duration, there are phonetically short consonants in the language, and there are geminates. The potential opposition is, however, not part of a quantity system, since this is a morphologically conditioned difference.

Thus we have to state that we found no contrastive use of duration in the language used by our Erzya speakers. The same applies to tone — there were no independent contrasts associated with pitch.

This means that duration and pitch are available for serving as phonetic cues to stressedness. Intensity (perceived subjective loudness) is the third feature usually marking a syllable as being stressed. Our study did not focus on intensity directly, since this would have demanded a laboratory setting that was not available at the time the recordings were made (for instance, maintaining a fixed distance between the speaker's mouth and the microphone), but we selected words where position of stress is estab-

lished by tradition. Since all words were produced in the same frame utterance, the influence of sentence rhythm was expected to be the same.

The results concerning stress are the most interesting — perhaps partly so because they are somewhat ambiguous and point toward directions that future research might take. Our results show that neither duration nor pitch serve as reliable stress cues. It turned out that vowel reduction is the feature that distinguished unstressed syllables from stressed syllables. Vowel reduction does not constitute substitution of a stressed vowel by a reduced vowel, but is a continuum, a process by which vowels can be gradually more or less reduced, probably depending at least partly on speaking style.

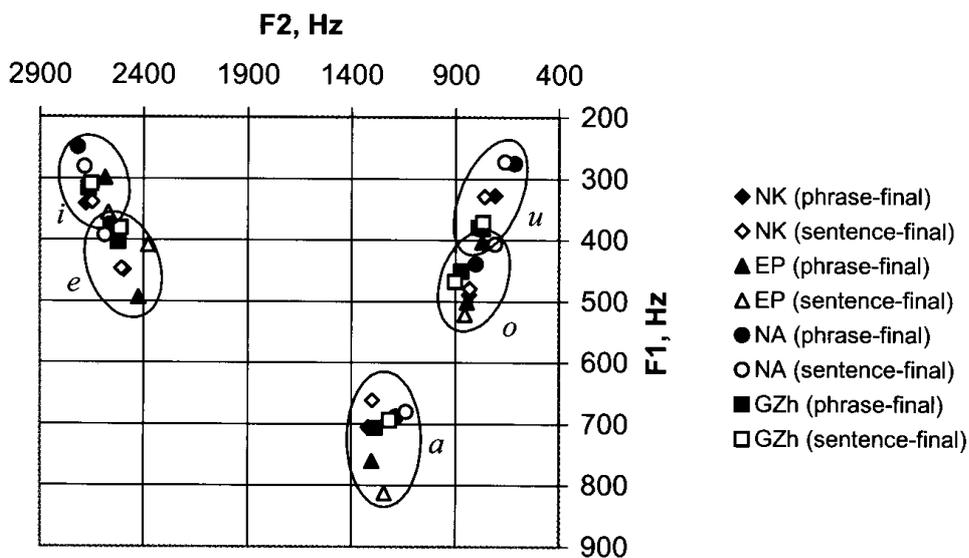


Figure 1. Acoustical vowel diagram of stressed vowels in the speech of four female informants. Speakers are identified by initials.

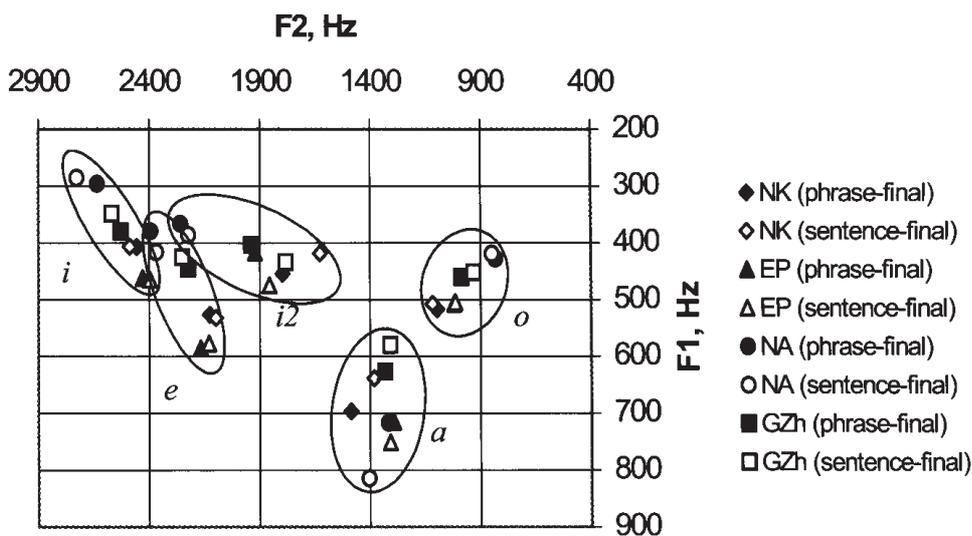


Figure 2. Acoustical vowel diagram of unstressed vowels in the speech of four female informants. Speakers are identified by initials.

Figures 1 and 2 offer acoustical vowel diagrams of stressed and unstressed vowels in the speech of the four female informants.<sup>2</sup> The speakers are identified by initials, and separate measurements are reported for phrase-final and sentence-final occurrences of the test words. Figure 1 represents stressed vowels, Figure 2 unstressed vowels. The stressed vowels occupy relatively extreme positions within the vowel space, while the unstressed vowels show varying degrees of centralization. There is hardly any overlap between the two distributions.

We did not use systematic listening tests, but had two listeners identify the position of stress, and as mentioned earlier, there was one set of words where their opinions differed. I believe the way we differed constitutes support for the role of vowel reduction in the perception of stress in Erzya.

Figure 3 shows average durations of vowels in trisyllabic words.<sup>3</sup> The left-hand set of four columns represents words in which both listeners agreed that it was the first syllable that was stressed. There are two values for the third syllable — the first shows the duration of the vowel when the third syllable was open, and the fourth shows the duration of the vowel in a closed final syllable.

On the right hand side are corresponding values for trisyllabic words where the author heard the second syllable as being stressed, but the native speaker heard the first syllable as being stressed. Evidently I had based my judgment on the greater duration of the second syllable, and the native listener insisted that it is the first syllable that is stressed, even though it contained the shortest vowel. (Neither one of us heard the third syllable as being stressed, regardless of the fact that it could have the longest vowel in the word).

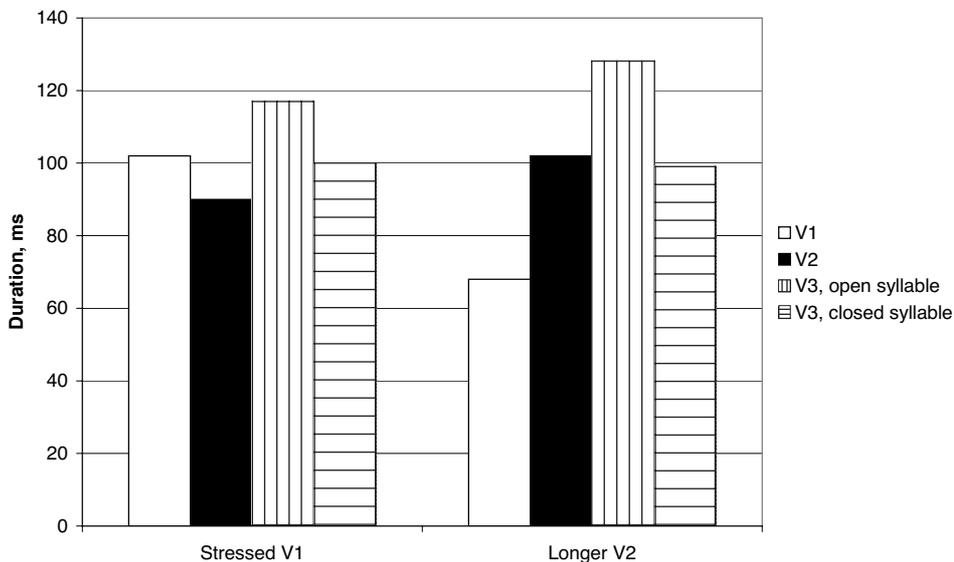


Figure 3. Vowel durations in trisyllabic words.

<sup>2</sup> Figures 1 and 2 replicate Figures 4 and 5, pp. 62 and 63 of "Erzya Prosody" (2003).

<sup>3</sup> Figure 3 is based on Table 3, page 54 of "Erzya Prosody" (2003).

Formant measurements showed that the native listener's judgment was based on unreduced vowel quality; the author's decision was evidently influenced by the durational structure of the test words. We agreed that systematic listening tests are called for, especially with synthesized test materials where each variable can be independently controlled.

### **3.1. Meadow Mari prosody: Introduction**

For some reason, the prosody of Meadow Mari has attracted the attention of several phonologists, who have offered differing theoretical interpretations of its structure.

Meadow Mari has full and reduced vowels; depending on the dialect, one or more reduced vowels are separate phonemes in the language, so that it is possible for a word to contain only reduced vowels, one of which will be stressed.

The terminology here is misleading. "Reduced vowel" is a term that normally refers to some kind of "full" vowel that has been reduced, has become more schwa-like. The so-called reduced vowels in Mari are central vowels, comparable in their acoustic structure to, for example, Estonian [õ]. They do have a special characteristic that sets them apart from the "full" vowels: they are shorter. The difference is not great, but it is statistically significant. And unstressed "reduced vowels" are likewise subject to vowel reduction, just like "full" vowels.

In some analyses the distinction of full and reduced vowels is presented as a principal phonological contrast (Kangasmaa-Minn 1998). In other analyses, an attempt has been made to define the alternation of full and reduced vowels as an opposition in vowel length (Hayes 1985).

Word stress has been claimed to be non-phonemic in Mari, but unlike in Erzya, it appears to be determined by the segmental structure of the word. A common claim is that stress falls on the last full vowel of the word; if there are no full vowels (i.e. if all vowels are reduced vowels), the initial syllable is said to receive stress. In a modern theory of phonology, namely Optimality Theory, the system is described as a DTO system (default to opposite side), where stress falls on the rightmost heavy syllable; in the absence of heavy syllables, stress falls on the leftmost syllable (Baković 2004).

It should be mentioned that in all former experimental studies the number of speakers has been very small — 1 to 3 maximum — and often it is unclear to what extent the results have been reliable enough for theoretical generalizations. A more thorough experimental approach thus appeared to us to be justified.

### **3.2. Meadow Mari prosody: Materials and method**

A test corpus was recorded from eight speakers with 100 test words, placed in a frame comparable to what we had used for the study of Erzya: "I said ..., not ... ." Every word occurred both in phrase-final and sentence-final position. The words consisted of one to four syllables, and were selected by T.-R. Viitso. Every speaker produced 200 test words, for a total of 1600 for the group.

Six of the speakers were born in Mari El, one in Bashkortostan, and one in the Perm region. Their birth years ranged from 1963 to 1981; seven of the eight were students either in Tallinn or at Tartu University (one was working in Estonia). Four of the speakers were recorded in 2000 in Tallinn at the Laboratory of Phonetics of the Institute of Cybernetics of the Tallinn Technical University; the other four were recorded in 2004 at the University of Tartu. The acoustic analyses were carried out at Tartu University, using the Kay Elemetrics CSL 4300B system and the Praat analysis program. The location of stress was ascertained by Vasilij Nikolajev, a native speaker of Mari, through repeated listening. Measurements were made by members of a team of graduate and doctoral students at Tartu University; the results were subjected to statistical treatment, as before.

The main research questions that were addressed in the analysis were the following:

- a) Is there contrastive quantity in Mari? What would be the linguistic status of possible phonetic differences in sound duration?
- b) What is the role of fundamental frequency in Mari prosody?
- c) What is the relationship between vowel quality and the prosodic system?
- d) What are the phonetic manifestations of stress?
- e) What is the role of stress in the phonological structure of Meadow Mari?

### **3.3. Meadow Mari prosody: Results**

Coming first to the role of duration, it appears certain that there is no contrastive quantity in the language. This conclusion is supported by the fact that duration has another function: it is the most reliable phonetic correlate of stress. There are geminate consonants in the language, but for several reasons, we have decided that this does not constitute evidence for contrastive quantity.

The reasons why we do not consider the existence of geminates to represent contrastive quantity are the following. There are intervocalic consonant clusters in the language, and the durational characteristics of geminates resemble those of consonant clusters. In numerous examples, the syllable boundary within a geminate is simultaneously a morpheme boundary. And in the sample that we analyzed, gemination did not appear to be systematic: it seemed to be restricted to only a small number of consonants rather than involving the whole consonant system. A final argument is typological: it is unlikely that a language has a quantity opposition in consonants while lacking it in vowels, whereas the opposite case is quite common in languages of the world.

We also found systematic preboundary lengthening in the materials that we analyzed — lengthening of the final syllable of the word in both phrase-final and in sentence-final position. An unstressed vowel in preboundary position was longer than a stressed vowel in other positions of the word.

The study also showed no contrastive use of fundamental frequency at the word level, which means that we found no indication of contrastive tone. The main function of F<sub>0</sub> at the word level appeared to be contributing to the identification of stressed syllables. While stressed syllables usually

had higher F0 than unstressed syllables, sentence intonation could override the F0 patterns associated with word stress.

Of the three prosodic features, duration and pitch appear to have no contrastive function in Meadow Mari. Stress, on the other hand, has a significant role, even though it is not totally unambiguous.

In our materials, stress position appeared not to be fixed with reference to a given syllable within a word. For example, in the set of disyllabic words, there were 29 words where the speakers stressed the first syllable, 13 words with stress on the second syllable, and in the case of 16 words, different speakers positioned stress differently. There was just one minimal pair in the corpus: /'šerge/, with stress on the first syllable, means 'dear', and /šer'ge/, with stress on the second syllable, means 'comb'. The words are spelled in the same way, and they were not presented as minimal pairs, so the readers were free to choose a meaning. Most speakers evidently chose the first meaning and produced the words with stress on the first syllable, but one female speaker produced the word with stress on the first syllable in the first occurrence, and with stress on the second syllable in the second occurrence. One male speaker stressed the second syllable both times.

As far as trisyllabic words go, 9 were produced with a stressed first syllable, 2 with a stressed second syllable, and 9 with a stressed third syllable. In seven words, there were differences between speakers: five words could be produced with stress on the first or third syllable, one with stress on the second or third syllable, and one word ([*ludanna*] 'we read', praet. II) could be produced with stress on the first, second, or third syllable. Now here the first and third syllable had full vowels, and the second syllable contained a reduced vowel, but this could be the vowel that received word-level stress.

The results of the analysis of duration and F0 may be summarized as follows. Duration is a reliable cue to the presence of word-level stress: a stressed syllable is longer than an unstressed syllable in the same position within a word. But this is not the only function of duration: it serves also to signal the presence of a boundary. In the analyzed materials, a final syllable of a word regularly had a lengthened vowel, regardless of the position of word-level stress.

The two functions of duration interact: a stressed final syllable is longer than an unstressed final syllable. However, in cases where word-level stress is on a non-final syllable, lengthening to signal the presence of the boundary was greater than the lengthening due to the presence of stress.

As was the case in Erzya, we also found that stressed syllables were characterized by relatively optimal vowel quality, i.e. lack of vowel reduction. Figure 4 is an acoustical vowel diagram of stressed and unstressed vowels, produced by four female speakers in phrase-final position.<sup>4</sup> Again, stressed vowels occupy extreme positions in the vowel space, while unstressed vowels show varying degrees of centralization.

I would like to point out in particular the position of the so-called reduced vowel. When it is stressed, its acoustic structure places it in the center of the acoustical vowel diagram — so it cannot become any more

---

<sup>4</sup> Figure 4 replicates Figure 17, page 76, of "Meadow Mari Prosody" (2005).

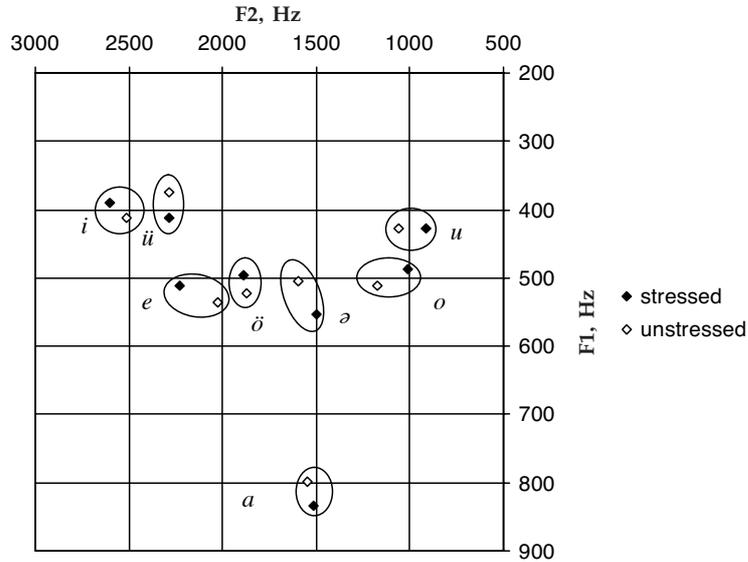


Figure 4. Acoustical vowel diagram of stressed and unstressed vowels occurring in phrase-final words produced by four female speakers.

centralized (if the process of vowel reduction is equated with vowel centralization). But while the other unstressed mid vowels move closer to the center, unstressed [õ] moves higher and a little toward the front. Thus optimal vowel quality does constitute another quite reliable cue to the presence of stress.

Some earlier analyses had claimed that if there were both full vowels and reduced vowels in the word, the reduced vowel was unstressed. We were not able to confirm this claim — our speakers did produce some words that had both full vowels and reduced vowels, and it was the reduced vowel that was stressed. But when all vowels in the word were reduced vowels, stress did occur primarily on the first syllable.

#### 4. Current project — Livonian prosody

The language we are currently working on is Livonian. Livonian occupies a special place among the Finnic branch of the Finno-Ugric languages. The Livs were the first inhabitants of the land to which they gave their name — Livonia, Livland, Liefland —, they were the first Finnic tribe to accept Christianity, which also meant subjugation by outside conquerors, which has led to the current very precarious status of the Livonian language. Contact with speakers of Baltic languages is usually believed to have resulted in a special feature of Livonian prosody: Livonian is the only Finno-Ugric language that has acquired or developed tone. As practically all current speakers of Livonian are bilingual in Livonian and Latvian, we are also looking at the language contact aspect. As the Livonian villages were located on the shores of the Baltic Sea, the language may also be considered a member of the linguistic community that is sometimes called Sprachbund, sometimes called a convergence area around the Baltic Sea.

There is a further reason to study Livonian, and that is the fact that Livonian is an endangered language. It is high time to record and analyze as much of its structure as possible. The study of Livonian prosody is part of the Finno-Ugric prosody project; we use the same methodology as with the other languages that we have already analyzed and that we hope to analyze in the future. We expect that the availability of comparable materials will also contribute to comparative studies in the future — which was part of the motivation for initiating the Finno-Ugric prosody project in the first place.

#### BIBLIOGRAPHY

- B a k o v i ć, E. 2004, Unbounded Stress and Factorial Typology. Optimality Theory in Phonology. A Reader, Malden—Oxford—Victoria, 202—214.  
H a y e s, B. 1985, A Metrical Theory of Stress Rules, New York—London.  
— 1995, Metrical Stress Theory. Principles and Case Studies, Chicago.  
K a n g a s m a a - M i n n, E. 1998, Mari. — The Uralic Languages, London—New York, 219—248.  
K e n s t o w i c z, M. 1994, Sonority-Driven Stress (Ms., Massachusetts Institute of Technology).  
U r b a n c z y k, S. 1999, Double Reduplications in Parallel. — The Prosody-Morphology Interface, Cambridge, 390—428.

*ИЛСЕ ЛЕХИСТЕ* (Колумбус)

#### ПРОЕКТ, ПОСВЯЩЕННЫЙ ПРОСОДИИ ФИННО-УГОРСКИХ ЯЗЫКОВ

Цель этой статьи — коротко познакомить читателей с проектом, над которым в течение последних шести-семи лет среди других работают и авторы представленных здесь исследований. Этот проект посвящен просодии финно-угорских языков.

К настоящему времени уже завершено изучение просодии эрзянского и лугово-марийского языков, в процессе работа над просодией ливского языка.