

PREFACE TO THE SPECIAL ISSUE OF ESUKA – JEFUL “ASPECTS OF SPEECH STUDIES”

Speech studies constitute a vast area of research, which includes phonetic and phonological theory, phonetic and phonological description of particular languages, the study of the acquisition and learning of phonetics and phonology, the study of phonological and phonetic variation and change, the study of speech production and perception, and of the paralinguistic aspects of speech, speech technology, etc. This special issue of the *Journal of Estonian and Finno-Ugric Linguistics* brings together studies representing different aspects of this large field of study.

One reason why speech is a complex object of study is that it is not only a linguistic phenomenon but also conveys non-linguistic information about physiological, psychological, social, and cultural factors. The study into this non-linguistic dimension of speech is represented in the present issue by the papers of **Rene Altrov and Hille Pajupuu**, and **Kairi Tamuri and Meelis Mihkla**. Both of these studies deal with the expression of emotions in speech. Tamuri and Mihkla test the modelling of emotional speech for the purposes of speech synthesis, building on earlier findings about the acoustic correlates of basic emotions (joy, anger, and sadness) in Estonian speech. Altrov and Pajupuu in turn show that the vocal expression of emotions is culture-specific and language-specific: test subjects representing different cultures and language families generally did not perceive the same emotions in Estonian speech as the Estonian listeners.

Cross-linguistic aspects are also central in the study of second language learning, given that one of the most important questions in this area concerns the influence of the learner's native language on the acquisition of a second language (L2). This line of research is pursued in the paper by **Einar Meister, Rena Nemoto, and Lya Meister**, who show that the production of L2 phonological features is influenced by the features of the learner's native language, although other factors, like orthography, also play a role. In particular, they find that the production of Estonian quantity contrasts by Japanese learners reveals the differences in the contrastive use of duration in the two languages.

Paralinguistic signals and cross-linguistic influence cause variation in language production, but variation may also reside in the phonology of a language. Phonological variation in Estonian is the topic of the paper by **Mari-Liis Kalvik and Liisi Piits**, who concentrate on the development of a suitable methodology for the study of phonological variation, and present the results of a pilot study on the extent and factors of the variation of three features: quantity contrast, palatalization, and the pronunciation of word-initial *h*. On the whole, they find variation to be more limited than could be expected on the basis of the existing literature. As the factors of this variation, they identify speech situation and the regional background of the speaker.

While Kalvik and Piits examine aspects of Estonian word phonology and prosody, the papers by **Heete Sahkai, Meelis Mihkla and Mari-Liis Kalvik**, and by **Heete Sahkai and Ann Veismann** contribute to the study of the phonology and phonetics of Estonian by describing aspects of Estonian sentence prosody, specifically, the realization and distribution of sentence stress. The first of these papers examines the acoustic correlates of the emphatic realization of the primary sentence stress, which serves to distinguish sentence-final narrow focus from broad focus. It identifies duration as the main correlate of emphasis, intensity as a secondary correlate, and the relative scaling of pitch accents within the utterance as a minor correlate. Sahkai and Veismann in turn find that the placement of the primary sentence stress in Estonian is determined, among other things, by predicate-argument structure. This finding confirms that Estonian belongs to the category of languages with “plastic” intonation.

A very important area in modern speech research is the applied field of speech technology, in particular, text-to-speech synthesis and speech recognition. All the papers in this issue have something to contribute to this area of research.

A central task in the current development of text-to-speech synthesis is to achieve as natural-sounding synthetic speech as possible. This involves both the modelling of paralinguistic aspects of speech, for instance the acoustic correlates of emotions, and of linguistic aspects, like segmental phonetics/phonology and sentence prosody. The study by **Tamuri and Mihkla** is directly devoted to this goal, as they propose acoustic models for three basic emotions, which can be applied in the synthesis of emotional speech; the results of **Altrov and Pajupuu** in turn show that the acoustic modelling of emotions is a language-specific task. The paper of **Kalvik and Piits** is likewise directly inspired by the

current needs of the development of Estonian speech synthesis, as issues of phonological variation must be resolved in the most natural way for the purposes of the conversion of written text into speech sounds. The studies by **Sahkai, Mihkla and Kalvik**, and **Sahkai and Veismann**, in turn, constitute a prerequisite for the modelling of sentence prosody in speech synthesis.

The development of speech recognition, too, must reckon with the natural variability of human speech. This includes speech produced by language learners, which is the topic of the study by **Meister, Nemoto, and Meister**.

Heete Sahkai, Meelis Mihkla, and Hille Pajupuu