https://doi.org/10.3176/lu.1979.3.09

Estonian Papers in Phonetics [ed. A. Eek], Tallinn 1972—77 (Institute of Language and Literature, Academy of Sciences of the Estonian S.S.R.)

1. Experimental phonetic research in Tallinn.

The decisive impetus for experimental phonetic studies in Estonia was given by Professor Paul Ariste, who studied first (1932) under the supervision of Professor F. Äimä in Helsinki and later (1932—33) in Hamburg under Professor G. Panconcelli-Calzia. We can find this piece of history in the third volume of the Estonian Papers in Phonetics (= EPP) which came out in 1974, and which is dedicated to the mentor of the present generation of Estonian phoneticians on the occasion of his seventieth birthday in 1975. In this volume P. Ariste writes about «Wie Professor G. Panconcelli-Calzia Lautgrenzen festgestellt hat» and G. Liiv gives a survey of Paul Ariste's scientific work.

The Experimental Phonetics Laboratory was founded at the Institute of Language and Literature of the Academy of Sciences of the Estonian S.S.R. in 1961. Since then intensive experimental research has been carried out by the staff members of the laboratory. The first issue of EPP was published in 1972, but several papers had already been published elsewhere especially by Georg Liiv, Arvo Eek, Mart Remmel and Kullo Vende. The laboratory was reorganized in 1977 as the Department of Computational Linguistics of the Institute of Language and Literature, Academy of Sciences of the Estonian S.S.R. After the first issue of the Estonian Papers in Phonetics in 1972 six further issues have been published. In the following survey the very interesting latest publication (1978), which is a collection consisting of papers read at a symposium dedicated to problems of accentology (Tallinn, November 1978), is not dealt with. The staff members of the Experimental Phonetics Laboratory have works published also elsewhere than in EPP. The bibliography of these works can be found on the final pages of the annual issues. Since 1974 there has been cooperation with the Institute of Cybernetics of the Academy of Sciences.

In the following survey the main attention is focused on topics which belong to linguistic phonetics. Other matters — among them research on low-frequency signals (Rohtla et al. 1974), stress manifestation in the Estonian folk song (Lippus 1977), and the theory of consonance (Ross 1977) — must unfortunately be left beyond the scope of this review.

2. Instrumental development.

The Experimental Phonetics Laboratory has been very active in developing new instrumental methods. The following methods are especially noteworthy:

1) Use of a 52-channel dynamic spectrograph which was available already as early as 1963. The method enables the recording of dynamic spectrograms and short time sections.

2) Spectrography by means of a computer using an A/D converter between speech and computer (see M. Remmel, A System for Digital Processing of Speech, 1972 and A. Eek, Acoustical Description of the Estonian Sonorant Types, 1972).

3) Measurement of fundamental frequency and overall intensity by means of a computer. The computer intonogram in which the fundamental frequency curve is observable in semitones and the overall intensity in dB values shows also the digital values of F_0 (semitones and Hz values) and intensity. The program was elaborated by M. Remmel (cf. K. Vende, Phonetic conditioning factors of pitch in Estonian vowels, 1973).

4) Dynamic palatography for registering linguopalatal contact and its changes during continuous speech (see A. Eek, R. Haavel, O. Künnap, M. Remmel, M. Veigel, A technique of dynamic palatography in application with a computer, 1973).

5) Electromyography for the study of the facial muscles using surface electrodes. M. Piirmets and H. Ristmets are responsible for the technical part of the method (see A. Eek, M. Piirmets, H. Ristmets, A preliminary electromyographic study of facial muscles in Estonian labial consonant production, 1976).

It can be pointed out, too, that X-ray

cinematography for the study of articulation was developed and applied before the papers in EPP. In several publications of EPP perceptual tests combined with elaborate statistical treatments are applied. In some papers synthetic stimuli are used. In the majority of cases when a computer has been used M. Remmel has been responsible for the design of the programs.

3. Problems related to qualitative sound segments.

3.1. In the paper «Acoustical Description of the Estonian Sonorant Types» (1972) A. Eek tries to find distinctive spectral properties for the Estonian sonorants */l, ɾ, m, n, ń, ŋ, r/* produced in isolation by three informants. It was difficult to reveal distinctive properties for nasals although they can be distinguished as a separate class from other sonorants. Eek concludes that "there are sufficient grounds for assuming that the excellent identifiability of the nasals articulated in a vowel context is to a great extent due to the differences in the terminal values of F2 of the vowel" (p. 23).

3.2. In the paper "A Pilot Study of Nasalization in Estonian" (1972) A. Eek gives a valuable list of the acoustical effects on vowels which are caused by coarticulatory nasalization in nasal environments. The environments with liquid consonants and those with nasal consonants are compared with each other.

3.3. The phonemic palatalization in Estonian has been examined by A. Eek in the paper "Observations in Estonian palatalization: an articulatory study" (1973). Dynamic palatography has been applied. Intervocalic consonants */l, ɾ, n, ń, s, ś, t, ʈ/* occurring in disyllabic words were read twice by the author. It was found that the vowel environments influenced the area of the lateral lingual contact, but less in palatalized than in non-palatalized consonants. The main results are as follows: 1) the *i*-like transition preceding a consonant and the initial part of the consonant itself remain the main bearers of Estonian palatalization, 2) the final part is entirely subjected to the coarticulatory influence of the following vowel (p. 31). In the Russian (ref. to studies by others), on the contrary, the palatalization is applied to the final part of the consonant and to the beginning of the following vowel. Very informative figures illustrate the discussion.

3.4. In the paper "A preliminary electromyographic study of facial muscles in Estonian labial consonant production" (1976) by A. Eek, M. Piirmets and H. Ristmets three muscles of the labial area have been examined: musculus orbicularis oris superior (OOS), m. orbicularis oris inferior (OOI) and m. depressor labii inferioris (DLI), i. e. the muscles which can be assumed to be responsible for labial articulation. Among other things the century-old geminate rearticulation hypothesis has been studied. It was found that OOS has functions in lip closing, closure preservation and lip rounding, but OOI has functions only in lip rounding and closing, not in closure preservation. It is claimed that OOI executes functions also in closure opening and lip retraction (p. 17). DLI is active in the opening movement. The relationship between the pressure behind the labial closure and the muscular effect needed for opening the closure (cf. stops and nasals) is discussed (p. 19). The rearticulation hypothesis is rejected, because it was found that two-peak cases in single consonants and one-peak cases in geminates were possible (p. 19). Gemination is interpreted as a perceptual phenomenon (p. 21).

4. Duration and quantity.

The papers "Observations on the duration of some word structures" I (1974) and II (1975) by A. Eek deal with the temporal organization in mono- and disyllabic words. The inherent consonant durations were measured. The theory of temporal compensation which assumes that duration is programmed on a longer speech unit than a sound is discussed intensively. The word has been regarded as an isochronic unit of a temporal program (by Shockey, Gregorski and Lehiste 1971), but it is suggested on the other hand "that we have reason to assume that the rhythmic foot is a unit of isochrony in longer words or utterances" (ref. to Lehiste 1973). The foot is defined as a sequence of a stressed syllable plus all the following unstressed ones. Almost a complete inversely proportional relationship between the intervocalic consonant carrying quantity contrast and the following vowel was observed (p. 24). The

initial consonant and the vowel of the primary-stressed syllable do not depend essentially on the quantity degree of the following consonant (p. 24). If Q1, Q2, Q3 consonants cover respectively 13, 23, 33% of the duration of the whole word, then the second-syllable vowels in the corresponding words cover 33, 22, 15% respectively. A strong inversely proportional relationship exists also between the quantity contrastive vowel of the primarystressed syllable and the second-syllable vowel in CV(V)CV words. The half-longness of the second-syllable vowel is discussed on p. 28. Eek concludes that in the case of small segmental durational differences of Q2 and Q3 the durational difference of the second-syllable vowel is an important cue in the perception of quantity degree. Note that Posti's two degree interpretation is partly based on the duration of the second syllable. The differences in temporal organization in isolated words and in sentences are mentioned (ref. to final lengthening observed by Lindblom in 1968).

In the second paper (1975) the quantity relationships of Estonian have been studied more exhaustively. The division of words up to ten-syllable long into mono-, di- and trisyllabic rhythmic units has been examined in greater detail. It is stated that the contrast of three quantity degrees is possible only in a stressed syllable; in an unstressed syllable quantity degree 3 is impossible. The exact identification of three quantity degrees presumes the existence of di- or trisyllabic rhythmic units. It is suggested that the quantity degree depends on the interrelations of three relatively autonomous factors — duration, stress and tenseness. Tenseness marks the particularly intensive articulation of the sound(s) of a stressed syllable (p. 28).

5. Pitch and intonation.

5.1. The first paper to deal with tonal features is K. Vende's article «Intrinsic Pitch of Estonian Vowels: Measurement and Perception» (1972). The aim is to show to what an extent the suggestion that higher vowels have a higher fundamental frequency holds in Estonian. This phenomenon was noticed in fact already by J. W. Black in 1949 concerning American English (Journal of Speech and Hearing Disorders 14 3, pp. 216—221). The problem is whether this kind of phenomena are essential also for perception and how the listener responds to vowels in which the fundamental frequency is not constant but changing. The short vowels are treated. After a most sophisticated analysis of the material Vende comes to the following conclusions: The vowel quality and the relative height of its pitch are connected also in Estonian. The height of the tongue is not the only factor explaining the phenomenon, but also the advancement of the tongue and (what is more remarkable) the rounding of the lips. The listening tests showed that in a vast majority of cases the auditors were able to decide with great certainty which of the two vowels presented was higher. The average pitch is perhaps the leading parameter for the perceptual choice. The listeners were not able to make any firm decisions about the movement of pitch in short vowels. Note that the test words of the type */pVkk/* were cut out from the frame sentence *Nimi pVkk on haruldane.* Later Vende repeats the same conclusion in «On people's ability to analyse pitch movement» (EPP 1977). Vende notes also that the order of the vowels, when presented in pairs to the listeners, plays also a significant, though unconscious, part in the listeners' minds when they evaluate the difference in vowel pitch (1972 : 101).

In the paper «Phonetic conditioning factors of pitch in Estonian vowels» (1973) K. Vende continues to study — again very thoroughly — the effect of various contextual factors on the fundamental frequency of the vowel. The remote target is to examine the intonation of Estonian questions and answers, but it is necessary to know also other factors than those caused by the communicative type of the sentence (p. 46). Several side effects are examined.

In the article «Intonation of question and answer in Estonian: I» (1975) K. Vende attacks the main problem, the prosodics of Estonian questions and answers, which was in fact touched upon in the previous article, too. Single words of up to three syllables and sentences of two monosyllabic words were examined, all pronounced as yes-or-no questions and as answers. Listening tests were used. The concept of vocal gesture has been dis-

cussed. The listening tests showed that many utterances intended as questions were also heard as such by most listeners. Most of the utterances intended as answers were perceived by a majority of listeners as answers (p. 93). Utterances that had obtained 80% or more "question responses" are called "good questions" and utterances perceived by 80% or more of listeners as answers are called "good answers". When these "good" utterances were used as material in a re-test three years later, it was found that all good questions and answers remained so in the re-test and their number had even increased (p. 94). Side factors — which are examined at length — can influence the perception of the sentence type.

5.2. In the article «Experiments on the perception of emphasis and question» (1974) R. Haavel raises — with reference to Chistovich 1972 — a very useful question: how could one decrease the amount of information necessary for the description of pitch curves without distorting the parameters interpreted by man as prosodic characteristics of a sentence? The problem is relevant on the basis of the fact that because of the limited volume of the short-term memory of man no complete description of the pitch contour of the sentence exists in his nervous system (p. 33). Haavel is especially interested in frequency and time parameters of the pitch curve.

The emphasis (prominence, logical stress; p. 34) and interrogativity of synthetic variants of the Russian sentence *mama myla Manyu* 'mother washed Manya' were examined. The procedure used for varying the pitch curve has been described elsewhere (Haavel 1973). The height, rate of frequency change and temporal location of the pitch peak were varied. The listeners (37 altogether) were expected to estimate the emphasis on every word in a stimulus, so that the sum total of the estimates for one stimulus was equal to 3 points (e. g. 0, 0, 3), and to decide whether the sentence was interrogative or not (0 or 1). It was found that the increase in pitch frequency is an essential characteristic when estimating the interrogativity of a sentence. The sensitivity of perception of changes in the location of extremum pitch varies at different positions of the curve (at word boundaries the rate of change in the estimates was higher than inside a word). Note that the test arrangement made it possible to locate the pitch peak at a word boundary or any other point within the word. Haavel concludes that the perception of emphasis is categorial and that the word is the segment to which the pitch curve is related, though the hypothesis still holds that a stressed syllable may be regarded as this segment especially in the case of the perception of the interrogativity of a sentence.

Haavel agrees with Fry that for the determination of stress it is not the magnitude of the change in pitch frequency, but the presence of the change that is of importance (p. 38). The same amount of increase in pitch leads to the increase of the estimates of emphasis according to the position of the word in a sentence: the same level of pitch causes fewer estimates of emphasis in the first word compared with the second and third words of the test sentence. The same level of pitch on the third word causes more responses for interrogativity than on the second word. The peak at the first word caused no evident perception of interrogativity. The article «Temporal characteristics of the pitch contour» (1974) continues a discussion of the same topic.

The concept of emphasis remains somewhat obscure: the dichotomy of rheme and contrast which are possible on the last word in Russian sentences is not dealt with.

5.3. Estonian word and sentence intonation is examined in two further articles: one by U. Lippus and M. Remmel («Some contributions to the study of Estonian word intonation», 1976) and the other by U. Lippus, E. Niit and M. Remmel («Further results on the intonation of Estonian: intermediate cases between word and sentence», 1977).

Lippus and Remmel are especially interested in the turning points of the pitch curve and their intermediate distances. They use correlation analysis to show (1) dependencies between the parameters measured and (2) between the parameters and the phonologically or communicatively definable factors. The latter class includes questions and answers. Mono-, di- and trisyllabic one-word sentences with the stress on the first syllable

were employed in the tests. Two speakers produced the utterances as answers and questions. The measurable pitch curve has been divided into parameters which describe the curve in the properties of height, turning point and time. These resemble partly those used by Vende in 1972. A part of the parameters are raw ones, i. e. direct measurements; another part of them are derived ones, i. e. calculated on the basis of raw data. The derived parameters can indicate quite complicated internal relationships between the raw parameters.

The variable syllable number of the test words gave rise to the normalization problem. This was resolved by examining the correlation coefficients found between the parameters and functional classes. The authors conclude that the questions are pronounced quicker than answers. In the monosyllabic words, however, the pitch rise to the peak is longer than in the other word types. This is considered to be a consequence of the fact that the monosyllabic word-sentences have no means of indicating interrogativity other than peak height and location. The height of the initial pitch peak (in monosyllabic words) and the pitch height later in the utterance (di- and trisyllabic utterances) seem to correlate strongly with interrogativity (r = about + 0.8 in all classes). A significant property for questions seems to be also the extent of insyllable pitch movements.

In the second paper 178 disyllabic one or two word sentences were used that had been pronounced as questions and answers by the same speakers as in the first study. The stress was possible on the first or on the second syllable or on both. Correlation calculus was used again. It was found — among other things — that the question function was coded into the beginning of the second syllable. With reference to the previous work the conclusion was drawn: the question is coded into the beginning of the last stressed syllable of the phrase (p. 59). The parameter concerned was t_6, the rise time for the pitch curve in the second syllable. The question also has a strong correlation with the fundamental frequency peak and the amount of the frequency change in the second syllable (p. 59—60). A strong correlation between stress and fundamental frequency was observed (p. 61). The extent of the frequency changes was also significant for stress.

6. Latent tonal contrasts?

Four of the papers in EPP deal with an exciting chapter of Estonian phonetics: the possible tonal contrasts. The supplements of the two previous articles discussed under 5.3 above, the article «Experiments on the perception of some word series in Estonian» (1977) by A. Eek and «Some preliminary data on lexical tonal oppositions in Estonian» (1977) by E. Helimski belong to this group.

6.1. Already in an earlier article (1975) G. Liiv and M. Remmel noticed that the temporal location of the pitch peak differs according to the quantity degree in CV(V)CV words: in the Q1 and Q2 words the pitch peak is placed at $^2/_3$ and in the Q3 words at $^1/_3$ of the first-syllable vowel duration. The pitch contour of a stressed syllable thus contains a tone rise and a tone fall segment. An unstressed second syllable is characterized by a steady fall segment. In all quantity degrees these tonal segments assume somewhat different time relationships. Remmel deals with the same problem of the short illative case in the supplement of the paper by Lippus and Remmel (1976). In the illative case forms the pitch curve occurs with two "hats" (peaks). Remmel refers to the occurrence of a stød-like effect, primarily in diphthongs. Two possible quantity models, M1 and M2, are discussed. M1 implies a system of three degrees. M2 assumes the existence of an additional class of the illative case. In an online computer game in which ten auditors gave their responses to the computer-synthesized stimuli with three keys corresponding to the genitive (Q2), partitive (Q3) and illative (Q4?) cases Remmel obtained somewhat vague results, however. Individual differences are possible.

Remmel discusses some earlier descriptions and theories about the stød phenomenon in Danish and Latvian as well as the problems of tonogenesis and tone in general. He makes the suggestion that the quantity system of Estonian is to be considered as being in the phase of serious transition. When answering the question,

whether it is the beginning or the end of a certain stage, Remmel concludes that the former possibility seems more acceptable.

6.2. A. Eek (1977) deals with the quantity and tone problems concerning paradigmatic oppositions. He puts the essence of the problem as follows: some authors are of the opinion that the number of actual oppositions is considerably larger than it was previously thought to be, while others discount the disputed oppositions as a fiction (p. 7). According to the former opinion the following forms are distinctive: *padu — patu — pattu* (partitive) — *pattu* (illative). It seems that differentation is possible even in the case of monosyllabic words of an identical phonemic composition: *puud* 'weight unit, nom. sg. or tree, nom. pl.' — *puud* 'tree, part. sg.'; *koid* 'moth, nom. pl.' — *koid* 'moth. part. sg.'. In those cases the vowel of the opposition's second member is felt to be accentuated (diphthong with a longer second component). Eek does not discuss the possibility of a latent syllable boundary between the vowel members.

Problematic word forms were put in sentences, the sentences were read by seven speakers, the test words were cut out from the sentences and identified by a variable number of auditors in four different listening tests. Duration and frequency measurements were performed. The listening tests showed that words like *püti* ('tub, genitive'), *pütti* (partitive) and *pütti* (illative) are perceived correctly by the auditors considerably more than randomly and that durational cues of the words can be shown to correlate with the identifications. The possible differences between the abessive case and *da*-infinitive forms (e. g. *võita* 'butter, abessive, sg.' — *võita* 'win, *da*-infinitive') are also examined. After a quite extensive discussion Eek comes to the conclusion that perhaps these data point to the fact that the language is on its way towards differentiating the homophones that have arisen in the course of its development, although on the other hand it is not clear whether they have ever been true homophones. The progress of tonogenesis is usually started by the monosyllabic structure of morphemes and the functional load arising from homophony.

6.8. E. Helimski (1977) concentrates only on the possible tonal differences in lexical word pairs. He refers to E. Polivanov who proposed in 1928 that the pitch outline of a word changes in its different paradigmatical forms. Helimski has noticed that certain Estonian homonyms are in fact not real homonyms, because having the same phonemic composition they nevertheless differ at the suprasegmental level, e. g. the word *koor* 'bark of a tree' has a lower pitch than *koor* 'choir'. Helimski has examined a material of nearly 100 nouns, adjectives, pronouns and numerals, mainly mono- or disyllabic words of Baltic-Finnic origin. Two native speakers of Estonian were involved to a greater extent in the study. No experimental data were available. Helimski classifies the words according to syllable number and quantity (diphthongs are also dealt with). Concerning the monosyllabic words which contain a long vowel (in Q3) or a diphthong Helimski finds four tonal types (high level, high with pitch downshift in the middle, low level, low with a downshift).

6.4. In the supplement of the article by Lippus et al. (1977) it is stated that, by and large, the Estonian language is treated by phonologists as an intermediate case between the tone and accent languages — there is almost unanimous agreement on the issue that the first and second degrees are in a quantitative opposition and the second and third degrees in an accentual opposition. Such a consensus of opinion is based, however, not on phonetic arguments, but on better agreement of such a description with the given concrete morphophonology and morphology. In terms of phonetics, it is possible to describe also the opposition of the first and second degrees as an accentual opposition and that of the second and third degrees as a quantitative one (p. 66).

When explaining the present-day tonal discrepancies Lippus et al. refer to the possibility that they may originate from former differences of a physiological nature (ref. to Hombert 1976). They claim — on the basis of measurements that language history is still reflected in the pitch frequencies of the final *e*. When a word represents the older form in which the final *h* is dropped (*mure*, *pere*, *vaie*), the fundamental frequency of the vowel is lower (by an average 10 to 12 Hz only), if it is

compared with the cases in which the final *k* is dropped after *e* (*mine*, *ole*). The cases in which no drop has happened (*sale*, *kare*) are located between the other two (average 150 Hz in the experiments). Physiological study has shown that where *k* and *h* exist in language there is a falling pitch before *h* and a rising pitch before *k* (p. 67). The relevance of these observations is discussed in more detail. Note that in the group with a former final *k* at least some word forms represent the imperative mood. Thus the comparison is not performed totally under the same conditions.

The concept of the prosodic system of a given language as a mixture of three different types of devices — quantity, stress and tone — is discussed in an interesting way (p. 67—69).

6.5. In her article "On the perception of tone contours with simple frequency changes" (1977) E. Niit touches upon the problem of Scandinavian contacts and the possible influence of the gravis accent of Scandinavian origin. The dialect of the Estonian islands has had a direct contact with Swedish due to long-standing Swedish settlements in North-West Estonia and, probably, e. g. on the islands of Hiiumaa, Vormsi and Ruhnu (p. 76).

7. Other topics.

Three articles by M. Remmel deal with some broad features of Estonian («Some general characteristics of Estonian speech: I—III», 1973; 1974; 1975). The long-term power density spectrum, distribution of pitch periods and perturbation of pitch as well as the taxonomic properties of the spectrum of Estonian speech are treated.

The problems of the design of an automatic phonetic transcription system is discussed by M. Remmel in «The problem of the relationship between the manner of coding and symbol inventory in phonetic transcription of sound classes» (1976). It does not become quite clear for the reader if the remote objective is really to try to construct a technical device for automatic transcription. Optimism as to the successful achievement of such a goal is not very well founded.

The article «Traces of phonetical parametres in phonostatistical distributions" (1977) by M. Remmel uses statistics to show some interactions in phonotactic compositions. The method might very well be valuable in future work on historical phonology and phonetics.

8. Final comment.

One feature seems to be common to the papers dealing with prosodics discussed above: a too great identity is assumed to exist between perceptual pitch (tone) and measurable fundamental frequency (F_0). Note that in the Finnish stressed syllable a rise-fall fundamental frequency curve is very usual, but it still represents an accent produced by means of the expiratory effort (subglottal pressure). A more prominent expiratory effort combined with the signalling of stress causes a higher F_0 level (stated e. g. by P. Ladefoged in «Three Areas of Experimental Phonetics», London 1967). Listeners seem to be able to distinguish between tone and force although these may exhibit equal fundamental contours. Mixtures of force and tone are possible, of course, but the problem is what mixture is represented by Estonian or different styles of Estonian. The use of semitones only is not sufficient because the original causal relationship between force and rising F_0 is not considered. Note that the possibilities of simulating physiological states by speech synthesis are in this respect very limited. Intensity curves also have considerable shortcomings for the study of this problem.

ANTTI IIVONEN (Helsinki)