DIPHTHONGS VERSUS VOWEL SEQUENCES IN ESTONIAN

ILSE LEHISTE

This paper deals with the problem of distinguishing between diphthongs and vowel sequences containing a word boundary. The language in which the phenomenon has been studied is Estonian. The paper is based on acoustic analysis of 500 short sentences produced by one informant (the author). It is to be understood that generalizations drawn from this limited material serve only to set up working hypotheses which will be tested by analysis of a larger corpus of utterances, produced by several additional speakers.

There are nine vowels and a large number of diphthongs in the language. Of the nine vowels, [*a e i u*] may occur in stressed as well as unstressed position in any syllable; they may also occur as first or second element of a diphthong. The following combinations of these four vowels were studied: the diphthongs [*iu*], [*ei*], [*ea*], [*eu*], [*ai*], [*ae*], [*au*], and [*ui*], and the vowel sequences [*i*] + [*u*], [*e*] + [*i*], [*e*] + [*a*], [*e*] + [*u*], [*a*] + [*i*], [*a*] + [*e*], [*a*] + [*u*], and [*u*] + [*i*]. It was hypothesized that in a diphthong, V_1 would be similar in phonetic quality to a stressed short vowel, and V_2 to an unstressed vowel occurring in the second syllable of a disyllabic word. It was further assumed that the diphthongal sequence would differ from a sequence of the same vowels containing a word boundary in the relative stress of the components: in the diphthong, V_1 would be stressed, V_2 unstressed, whereas in the sequence containing a boundary, V_1 would be unstressed and V_2 stressed.

Broad-band spectrograms were made of all utterances. The sentences were also processed through an intensity meter and pitch meter (produced by B. Frøkjaer—Jensen, Engineering Firm of Electronic Instruments, Copenhagen) and displayed on a four-channel Mingograph (Mingograph No. 42—EM/122, Elema-Schönander, Stockholm). The results of the study are summarized in a series of acoustical vowel diagrams which were constructed on the basis of averages calculated from measurements made of broad-band spectrograms. Corresponding tables are likewise presented.

The first figure shows $F_1 - F_2$ positions of the stressed short vowels [*i e a u*] and positions of the same vowels forming the first element of a diphthong. The positions of stressed long and overlong vowels have likewise been plotted on the diagram. The figure shows that the phonetic values of stressed short vowels are remarkably similar to the first components of diphthongs. Both differ markedly from long and overlong vowels, which are phonetically close to each other.

Figure 2 compares the second components of diphthongs with target positions of stressed monophthongs. In the case of [*i*] and [*u*], the second components of overlong diphthongs are phonetically similar to long vowels; but with [*a*] and especially

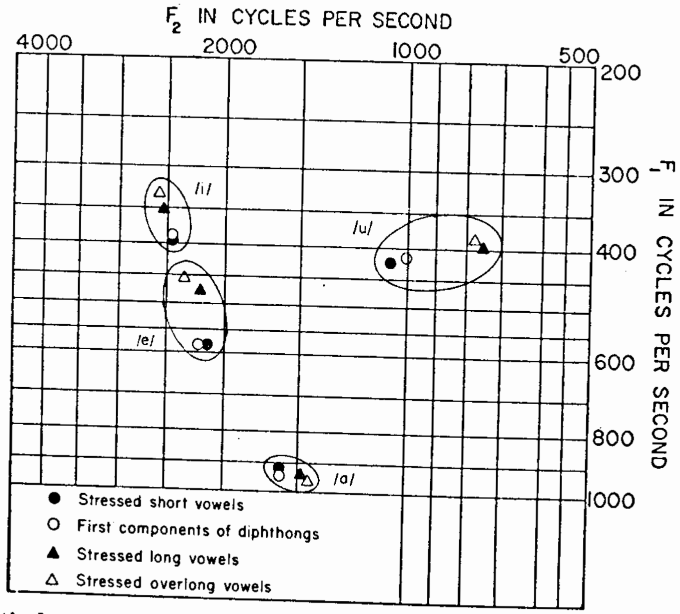

Figure 1. Acoustical vowel diagram representing $F_1 - F_2$ positions of the stressed vowels [*i e a u*] occurring in three quantities as monophthongs and as first components of diphthongs.

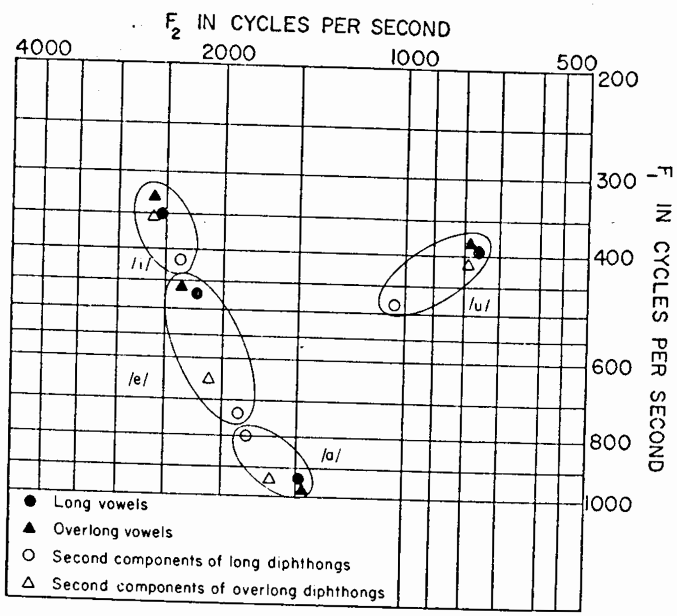

Figure 2. Acoustical vowel diagram representing $F_1 - F_2$ positions of long and overlong [*i e a u*] compared with the same vowels occurring as second components of long and overlong diphthongs.

with [*e*], no such statement can be made. The second components of long diphthongs are in each case distinctly different from stressed monophthongs.

Figure 3 compares V_2 of diphthongs with unstressed vowels in the second syllable of disyllabic words. The results are contradictory: for [*u*], unstressed vowels are similar in phonetic quality to V_2 of long diphthongs, but for [*e*], [*a*], and [*i*], the unstressed vowels are more similar to V_2 of overlong diphthongs.

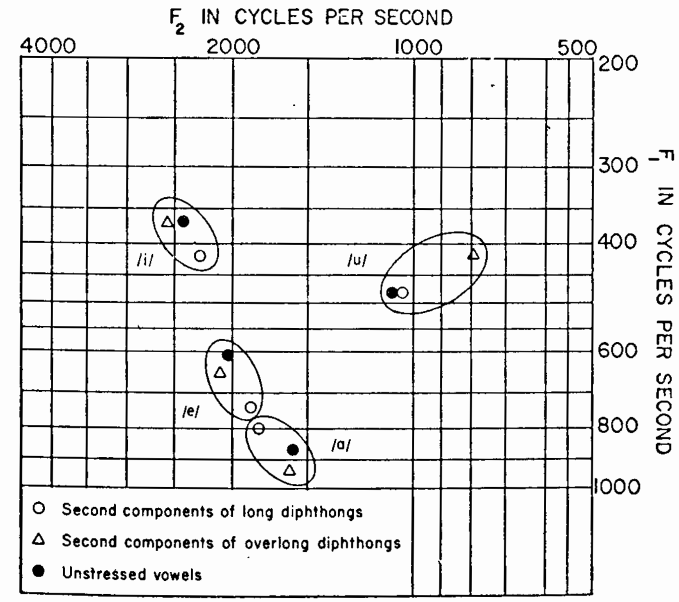

Figure 3. Acoustical vowel diagram representing $F_1 - F_2$ positions of [*i e a u*] occurring as second components of long and overlong diphthongs and as unstressed vowels in the second syllable of disyllabic words.

It seems that the second hypothesis cannot be verified on the basis of this material: the phonetic quality of the second component of diphthongs is not identical with the phonetic quality of unstressed vowels. Neither is it identical with that of stressed vowels, except for [*i*] and [*u*], where V_2 of overlong diphthongs was similar in quality to long and overlong monophthongal [*i*] and [*u*].

The clues provided by duration were considered next. The duration of the two segments was measured, assuming a boundary in the middle of the transition from the first to the second component. The two components of a diphthong were found to be almost equal in duration both in long and in overlong diphthongs: the lengthening of an overlong diphthong was apparently accomplished by a proportional lengthening of both components. The average durations of V_1 and V_2 of a long diphthong were 9.4 and 7.4 csec; in an overlong diphthong, the values were 13.2 and 13.9 csec. The average durations of vowels in a sequence containing a word boundary were 8.6 and 8.3 csec. Duration thus distinguishes a $V + V$ sequence from overlong diphthongs, but not from long diphthongs. The contribution of intensity to the difference between diphthongs and vowel sequences containing a word boundary came next under consideration. Only qualitative observations are available, since the test material was not suitable for a quantitative treatment of the data. Limitations of space make it impossible to present more than a single example.

Figure 4 contains an oscillogram, a fundamental frequency curve, and an intensity curve of the two utterances *Ilmad on kuivad* ("The days are dry") and *Seleks kulus terve eluiga* ("It took a whole lifetime"). The comparison is between the long diphthong [*ui*] in *kuivad* and the [*u*] + [*i*] sequence in *eluiga*, which contains a word boundary between [*u*] and [*i*]. The intensity curve shows an additional peak during the second component of the [*u*] + [*i*] sequence *eluiga*. Such separate peaks were frequently found in $V + V$ sequences; they did not occur in diphthongs.

In summary, a complex set of distinctions was found between diphthongs and $V + V$ sequences. The difference between long diphthongs and $V + V$ sequences was partly due to the difference in phonetic quality between the second component of diphthongs and between stressed vowels. The difference between $V + V$ sequences and overlong diphthongs was primarily durational. Intensity differences provided an additional clue that seemed to be relatively independent of vowel quality. An

Table I. Comparison of stressed monophthongs with the first component of diphthongs.

Vowel	Number of occurrences	Average Duration in csec	Formant frequencies in cycles per sec.		
			F_1	F_2	F_3
Short [*i*]	15	8.6	395	2495	3115
Long [*i*]	4	14.7	350	2540	3200
Overlong [*i*]	12	19.4	325	2560	3320
[*i*] as V_1*	2/3	11.5/14.0	380	2475	3035
Short [*e*]	13	9.6	585	2130	2940
Long [*e*]	7	12.1	470	2205	2885
Overlong [*e*]	12	18.3	460	2350	2985
[*e*] as V_1	11/16	8.4/12.9	580	2190	2975
Short [*a*]	21	11.4	925	1600	2695
Long [*a*]	8	15.9	950	1495	2720
Overlong [*a*]	23	22.1	975	1445	2765
[*a*] as V_1	25/18	10.2/13.9	955	1605	2705
Short [*u*]	10	9.7	420	1060	2780
Long [*u*]	14	17.2	395	750	2855
Overlong [*u*]	26	21.0	385	770	2845
[*u*] as V_1	5/4	6.8/11.3	410	1000	2730

* The formant positions for the first component of a diphthong are averages for both long and overlong diphthongs. The number of each type and the durations are given separately; the first number refers to the first segment of long diphthongs, the second to the first segment of overlong diphthongs.

Table II. Second components of diphthongs.

Vowel	Second component of long diphthongs					Second component of overlong diphthongs				
	N	Dur.	F_1	F_2	F_3	N	Dur.	F_1	F_2	F_3
[*i*]	28	6.4	415	2310	3055	19	12.2	365	2560	3215
[*e*]	2	6.0	740	1865	2800	8	13.4	655	2060	2875
[*a*]	2	8.0	800	1815	2775	7	13.4	950	1620	2700
[*u*]	11	7.5	480	1030	2725	7	16.0	415	785	2845

Table III. Unstressed vowels in the second syllable of disyllabic words.

Vowel	Quantity of preceding syllable	Number of occurrences	Duration in csec	Formant positions in cycles per second		
				F_1	F_2	F_3
[*i*]	Short	21	10.9	355	2520	3120
	Long	14	9.2	365	2470	3120
	Overlong	29	9.9	370	2465	3115
				Average: 365	2485	3120
[*e*]	Short	21	9.6	625	2030	2890
	Long	21	8.1	595	1980	2845
	Overlong	35	9.4	615	2005	2885
				Average: 610	2005	2875
[*a*]	Short	33	12.3	900	1530	2685
	Long	60	9.6	825	1615	2720
	Overlong	51	9.5	880	1625	2775
				Average: 870	1590	2725
[*u*]	Short	14	11.9	445	975	2810
	Long	9	9.7	505	1110	2815
	Overlong	21	8.9	470	1125	2735
				Average: 475	1070	2785

unexpected result was the discovery that the overlength of long diphthongs is distributed evenly over both components of the diphthong. This finding supports the assumption that the domain of overlength is the whole syllable rather than one of the segments of the diphthong.

DISCUSSION

Hint:

In the phonological pattern of Estonian the long vowels function as equivalent to long diphthongs and overlong vowels to overlong diphthongs. It is very interesting to hear now that they are similar in respect to their phonetic structure too, so far as quantity and duration are concerned. This conforms to phonological expectations.

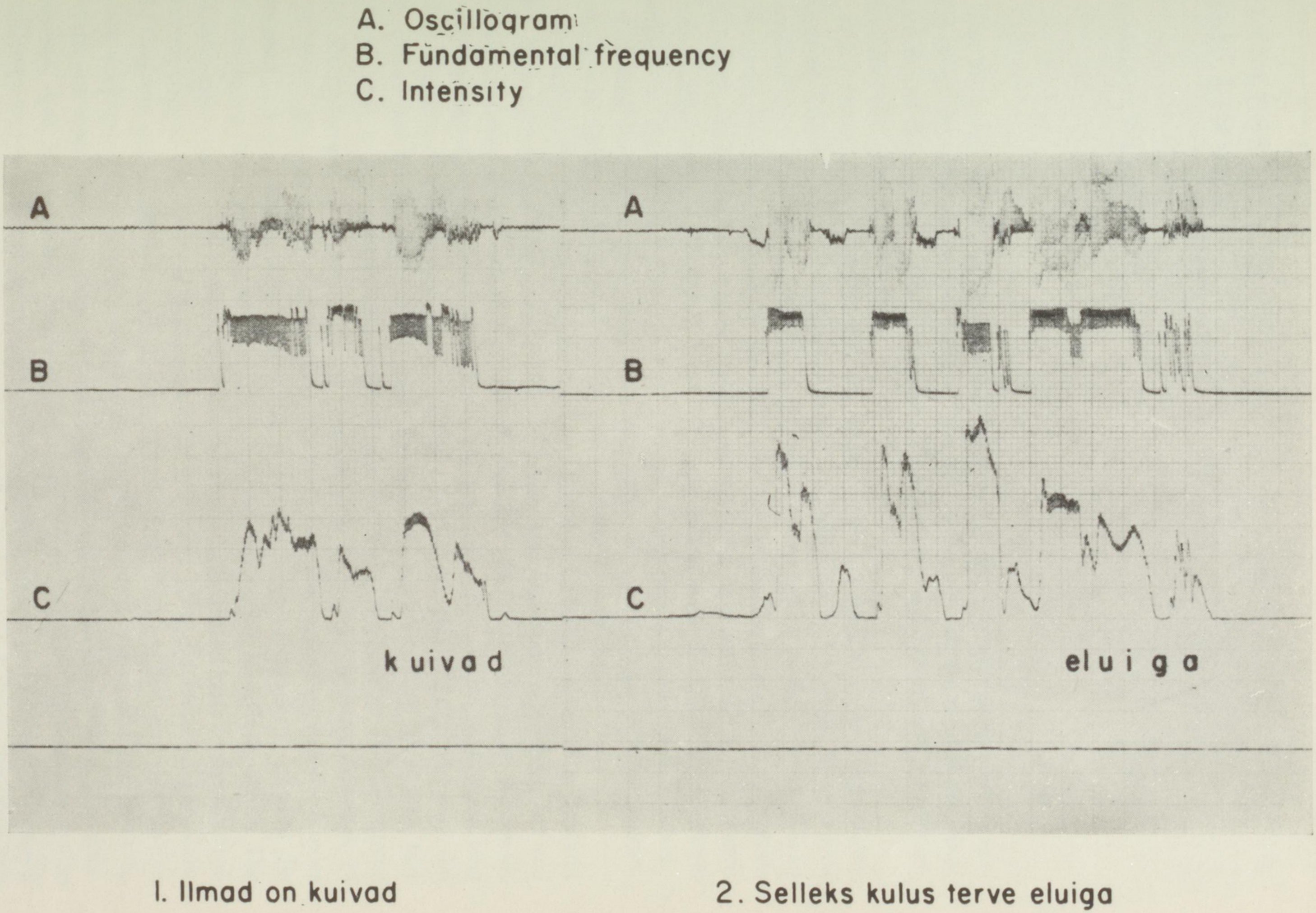

Figure 4. Oscillogram, fundamental frequency curve and intensity curve of two utterances, one containing the long diphthong [*ui*], the other the sequence [*u*] + [*i*].