

ROBERT HAMMARBERG (West Lafayette, Indiana)

SPECULATIONS ON THE ESTONIAN *e*-DIPHTHONGS

There is something suspect about the Estonian diphthongs which traditionally (in the orthography as well as in phonemic transcription) are represented as consisting of a vowel followed by an *e*. For one thing, no such diphthongs can be found mentioned in the current standard works on linguistic phonetics. This could, of course, be due to oversight — perhaps diphthongs on *e* are very rare. But the roots of suspicion go deeper than that, for it appears that the standard theory of linguistic phonetics does not allow for such things as falling diphthongs on *e*. Thus we are faced with a crisis: either the standard theory is wrong, or else Estonian does not really have diphthongs on *e*.

Diphthongs are notoriously troublesome from a taxonomic point of view. First of all, are we to regard them as single entities, as would be implied by the singular term “diphthong”, or is this merely a theoretically void cover term applied to intra-syllabic sequences of certain types of segments? Both views have their adherents. And, as in all matters of segmentation, the question cannot be settled on observational grounds, but must be decided on the basis of a complex of higher-level linguistic considerations. But whatever one’s views might be on the issue of segmentation, there is common agreement as to the fact that diphthongs do not end in the same way as they started. Thus the adherents of the mono-segmental view of diphthongs have described them variously as vowels with a changing formant structure¹, vowels in which there is an appreciable change of quality², as compound vowels³, and so forth. This view is obviously very uneconomical in comparison with the bisegmental view, according to which “diphthong” is simply a cover term for sequences of non-consonantal segments, and thus, in principle, no different from consonant clusters. But the bi-segmental view leads to another taxonomic problem. One of the two segments is going to be a vowel, constituting the peak of the syllable. But what is the nature of the other segment in the diphthong?

Chomsky and Halle define this other segment as a glide, and glides, in turn, are defined as being non-consonantal and non-vocalic. It is the specification for vocalicness that is of particular interest to us:

¹ B. Malmberg, *Structural Linguistics and Human Communication*, Berlin 1967, p. 52.

² H. A. Gleason, *An Introduction to Descriptive Linguistics*, New York 1961, p. 254.

³ E. R. Moses, Jr., *Phonetics: History and Interpretation*, Englewood Cliffs, N. J., 1964, p. 49.

Vocalic sounds are produced with an oral cavity in which the most radical constriction does not exceed that found in the high vowels [i] and [u] and with vocal cords that are positioned so as to allow spontaneous voicing; in producing nonvocalic sounds one or both of these conditions are not satisfied.⁴

Since it is obvious that both elements of a diphthong are sonorants (i. e., not obstruents), it is the constriction criterion that will be crucial for glides. Chomsky and Halle do not make clear what kinds of constriction they have in mind. A constriction by the tongue against the palate and velum is clearly implied by the reference to the high vowels, but, conceivably, labial constrictions as well as constrictions by the back of the tongue against the pharyngeal wall would also be possible. But under these criteria it would be impossible to produce a glide with the articulatory features of an unrounded mid vowel, such as implied by the Estonian $V + e$. There simply is no phonological parameter on which an *e* could achieve any kind of constriction.

There are many problems with this definition. For one, Chomsky and Halle apparently violate their own precepts by postulating a mid front glide, /*e*/, in the underlying forms of English.⁵ For another, this definition makes it impossible to account for rising diphthongs such as, e. g., Finnish *ie*, *üö*, *uo*, where the first element clearly is the syllabic peak, and the second therefore would have to be a glide. For reasons unrelated to be problems just mentioned, Chomsky and Halle later suggest substituting the feature syllabic for vocalic. This move perhaps avoids one problem by allowing for a perceptual definition of the entities involved (syllabic peak vs. non-peak), but Chomsky and Halle introduce another problem by claiming that

When vowels become nonsyllabic, they turn into glides: high vowels turn into the high glides [w] and [y]; nonhigh vowels into the nonhigh glides symbolized by [h].⁶

We need only note that Finnish has the contrastive intrasyllabic combinations *ih*, *üh*, *uh*, *ie*, *ia*, *üö*, *uo*, to demonstrate the inadequacy of this contention.

Up to this point, all suspicions have fallen on our taxonomic criteria, and rightly so — they are clearly inadequate. But the inadequacies, I believe, stem chiefly from the failure to take into account rising diphthongs of the *ie*, *uo*, etc. type, where the syllabic peak is on the first, less sonorous element. That is to say, the Chomsky-Halle criteria may, on the whole, be appropriate to account for falling diphthongs. Let us proceed on this assumption.

A falling diphthong, then, is to be regarded as consisting of a vowel plus a glide. The glide, in turn, is to be manifested as an articulatory constriction. The crucial thing to consider now, is the degree of this articulatory constriction. The characterization given by Chomsky and Halle, i. e., that the constriction be equal to or greater than that found in the high vowels (but not so great as to defeat spontaneous voicing), must be taken to be the canonical form of the various glides. In context, the canonical form may be altered, i. e., there are going to be allophones of the glide. The degree of constriction attributed to the canonical form would be found only in position before a tautosyllabic vowel. In post-vocalic (but still tautosyllabic) position, however, the actual degree of constriction attained would not be crucial. Consider the following description offered by Westermann and Ward⁷.

⁴ N. Chomsky, M. Halle, *The Sound Pattern of English*, New York 1968, p. 302.

⁵ N. Chomsky, M. Halle, *op. cit.*, p. 176.

⁶ *Ibid.*, p. 354.

In a falling diphthong it is possible to state with considerable accuracy where the tongue starts and the direction in which it glides, but not exactly how far it goes... It is customary to write them with two vowel letters, the first representing the starting-point of the glide and the second the direction in which the tongue moves: e. g. for the English diphthong *ai* in *time*, the tongue starts from near to the Cardinal 4 position and glides towards (but does not reach) the *i* position.

The crucial point here is that the glide element of the diphthong is realized as a movement in the direction of a point which is not reached. If direction of movement, rather than exact point of termination, is the defining characteristic of the glide portion of diphthongs, then there could hardly be a distinction between diphthongs on *i* and diphthongs on *e*. Starting out from a low vowel, the direction towards *i* and *e* would be the same, and to claim a distinction between the two would amount to claiming that exact distance of movement would be crucial. Yet it is well known, that in a diphthong on *i*, an actual *i* position (high and front) is not reached. This can be easily ascertained by recording an *ai* diphthong and then playing it backwards; it will not sound like *ja*.

The case against diphthongs on *e* can be summed up as follows: Falling diphthongs consist of vowel plus glide, and the glide is realized as a movement toward, but not to, one of the extreme points of the vowel space where an articulatory constriction is possible, a constriction which is greater than that found in any vowel, but less than that found in any obstruent. There are three points in the vowel space where the tongue could achieve such constriction: the high front area, the high back area, and, at least in theory, in the low back area. But no constriction of the required kind could occur in the mid front area implied by the assumption of diphthongs on *e*.

What then could the glide element of these *e*-diphthongs be? It is clearly not low and back, and it is equally clearly not rounded. By elimination we arrive at the conclusion that it must be high and unrounded. But if it is in contrast with non-syllabic *i*, then the only free spot left would be that of non-syllabic *i*, i. e., a non-syllabic high back unrounded vowel. On general grounds this would be a rather satisfying conclusion. We are assuming that the post-vocalic glides appearing in diphthongs are allophones of the glides which may occur in antevocalic position. Ladefoged⁸ mentions three such glides in the high region: the palatal (high front unrounded) /j/, the labial-palatal (high front rounded) /y/, and the labial-velar (high back rounded) /w/, representing, respectively, the nonsyllabic /i/, /ü/ and /u/. There is a gaping hole in the (non-labial) velar slot, which could be filled by that Estonian "e". This would bring about the kind of symmetry in the universal inventory of segments that is a characteristic of correct taxonomic systems.

Admittedly the argumentation that has led us to this point has been speculative and is based on assumptions that may well be vulnerable. Nevertheless we can find corroborating evidence for the conclusion reached, which ought to prevent it from being dismissed out of hand. Such evidence is to be found in the historical origin of these *e*-diphthongs. Kettunen gives two sources of these diphthongs. One source would be the Proto-Balto-Finnic (PBF) *i*-diphthongs, via a lowering process which occurred regularly before /v/ and /r/, but also in other environments.⁹ But there is nothing about a /v/ or an /r/ that would in any way explain

⁷ D. Westermann, I. C. Ward, *Practical Phonetics for Students of African Languages*, London 1933, p. 44.

⁸ P. Ladefoged, *Preliminaries to Linguistic Phonetics*, Chicago 1971, p. 60.

⁹ L. Kettunen, *Eestin kielen äännehistoria*, Helsinki 1962, p. 140.

why such a lowering would take place. These instances, therefore, provide us with no clues as to the nature of the *e*. The other source of *e*, however, will prove more illuminating. This is the instance where the *e* has developed out of a velar stop which in turn was followed by a liquid. There are about a dozen words which in all likelihood contained *Vkr* or *Vkl* sequences in PBF, and which now appear with *Ver* or *Vel* in modern Standard Estonian. The interesting fact is, that the cognates of these words in modern Standard Finnish appear with back rounded /*w*/-diphthongs. Moreover, in several present-day Estonian dialects we find these same words with /*w*/- or /*j*/-diphthongs, as well as, in a few cases, simply a lengthened vowel. Thus, for example ¹⁰:

| | | | | |
|----------|----------------|----------------|----------------|----------------|
| PBF | * <i>kakra</i> | * <i>kakla</i> | * <i>pakla</i> | * <i>nakla</i> |
| Standard | | | | |
| Finnish | <i>kaura</i> | <i>kaula</i> | <i>paula</i> | <i>naula</i> |
| Standard | | | | |
| Estonian | <i>kaer</i> | <i>kael</i> | <i>pael</i> | <i>nael</i> |
| Estonian | | | <i>paul</i> | <i>naul</i> |
| Dialects | <i>kaar</i> | | <i>paal</i> | <i>nail</i> |

Since the glide elements of these diphthongs are assumed to have arisen from velar stops, let us take a look at what would happen if such a stop first became voiced and continuant (as is generally assumed to have happened), and then lost its obstruent properties. The first step, that of /*k*/ → /*ɣ*/, would have yielded a segment with the following properties (among others):

| |
|---|
| — vocalic + consonantal — sonorant + high + back — low + voice + continuant — round |
|---|

The step changing this /*ɣ*/ to a glide would involve only altering the values for the features consonantal and sonorant:

| |
|--|
| — vocalic — consonantal + sonorant + high + back — low — round . . |
|--|

This is exactly the high back unrounded glide /*j̥*/, the nonsyllabic *i̥*, that we have been discussing.

This leads me to suggest the following scenario. The first step in the “deconsonantalization” of the original PBF velar obstruent would have led to the appearance of a velar, i.e., high back unrounded glide. This glide has been preserved in some dialects of present-day Estonian, among them the dialect making up the standard language. This would be

¹⁰ These examples are taken from the SKES.

the segment which the traditional orthography represents with the letter *e*.

At one stage, certain dialects of PBF would then have had a full complement of high glides: /j/, /y/, /j̥/, and /w/. The /j̥/, however, would have been highly marked, and thus very likely to change toward a less marked state. The marked features of /j̥/ would be the features [back] and [round], which would have the values + and —, respectively. A change in one or the other would yield a minimally marked glide. By changing the value of the feature [round] to + we would get a /w/. This is what Finnish and some of the Estonian dialects have done. If we change the value of [back] to —, we would get a /j/, which we find in a few cases. The change to /w/ is, however, the one most likely to occur, since it appears that the feature [back] dominates the feature [round]. That is to say, the value for rounding is more likely to be dependent on the value for backness than vice versa.¹¹

One further point needs to be made. If we are going to deny the existence of *e*-diphthongs in Estonian, then what about the *o*-diphthongs, *ao* and *eo*, which have been taken to be the mid back counterparts to the *e*-diphthongs? There appears to be some vacillation in the pronunciation of these sequences; both diphthongal and bi-syllabic pronunciation is found.¹² But, since diphthongal pronunciation does occur, must we allow for mid glides in the back region? In the case of the rounded glides, tongue height does not appear to be of crucial importance. Tongue position will determine whether the glide is front or back, but as far as the crucial constrictive movement is concerned, this will be provided by the lips. And here there are two kinds of labial constriction possible, one with lip protrusion, and the other with a narrowing of the lips, but without protrusion. The former could be the kind found in the *o*-diphthongs, the latter in the *u*-diphthongs.¹³

РОБЕРТ ХАММАРБЕРГ (Вест-Лафайет, Индиана)

РАЗМЫЛЕНИЯ ОБ ЭСТОНСКИХ ДИФТОНГАХ НА *e*

Определение нисходящих дифтонгов, представляемое обычно в работах по лингвистической фонетике, допускает только глайды, образующиеся при сравнительно высокой степени ротового сужения. Это исключает существование глайдов с невысоким или низким положением языка, таких как эстонские дифтонги, оканчивающиеся на *e*.

Если определение корректно, необходимо переопределить эстонские дифтонги на *e*. Оказывается, что единственным глайдовым элементом, который можно учитывать, является высокий негубной глайд заднего ряда, неслоговый *i̥*. Этот вывод подтверждает в определенной мере то обстоятельство, что многие из дифтонгов на *e* (которым в финском языке соответствуют дифтонги на *u*) исторически образованы от велярных смычных. Лишение велярного смычного его консонантных свойств как раз и ведет к появлению велярного (заднего и высокого) негубного глайда.

¹¹ N. Chomsky, M. Halle, *op. cit.*, p. 410.

¹² Cf. A. Raun, A. Saareste, Introduction to Estonian Linguistics, Wiesbaden 1965, p. 16.

¹³ This is the concept of in-rounding and out-rounding which has been applied to the vowels of Swedish. Cf. O. Jespersen, *Fonetik*, København 1899, p. 192; also B. Malmberg, *Distinctive Features of Swedish Vowels*. — For Roman Jakobson, The Hague 1956, pp. 316—321.