A PHONETIC EXPLANATION TO REDUCED VOWEL HARMONY SYSTEMS*

Mona Lindau

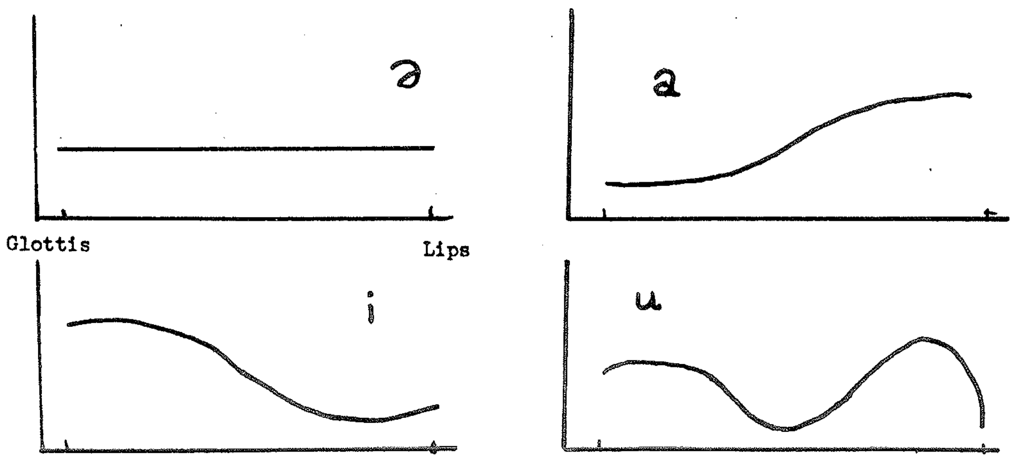
Proto-Kwa and other Proto-Niger-Congo languages are being reconstructed with a vowel system of five tongue-root advanced vowels and five tongue-root-retracted vowels (Stewart, 1971). There are modern Niger-Congo languages that still exhibit a ten vowel system, but most of them are reduced to nine- or seven-vowel systems, and in the case of most Lower Niger languages to eight-vowel systems. In this paper I will consider some common patterns of reduction, and attempt to provide a phonetic explanation for these patterns. The explanations involve predictions from a theory that is developed independent from theories of vowel production and phonological constraints on vowel systems, namely acoustic perturbation theory. Because of the independence of perturbation theory and theories of vowel systems the proposed explanation would be a theoretically strong one, provided it also stands up to closer scrutiny.

Perturbation theory[1] has recently been applied to theories of speech production to answer questions about the relationships between articulatory configurations and corresponding formant frequencies. We can ask questions about these relationships from two angles, either given a certain articulatory configuration and an articulatory change, what is the acoustic effect? or given a certain point in a formant space, and a certain formant change, what articulatory configurations could have accomplished this? At the moment the first question has had the larger amount of research devoted to it, so the discussion here will be restricted to questions of the first type.

---

* I wish to thank Kay Williamson and Peter Ladefoged for discussions and comments on earlier versions of this paper.

In applying perturbation theory to vowel production, the articulatory configurations are described in terms of area functions of the vocal tract. An area function "described the cross-sectional area of the oto-pharyngal cavity as measured perpendicularly to the longitudinal midline of this cavity. This midline runs from the glottis to the labial orofice of the mouth." (Öhman, 1973.) The cross-sectional area at any point along the vocal tract is calculated from the saggital distance at the same point along the vocal tract. Area functions of [ə, i, a, u] are illustrated below (Öhman, 1973, adapted from Fant, 1960).



The abscissa shows the distance from the glottis, the ordinate the cross-sectional area. The curves show approximate cross-sectional areas along the midline from glottis to lips. The vowel [ə] is simplified to a straight line, because the formant frequencies of this simple tube, closed at one end and open at the other, are very regularly distributed. F_1, F_2, and F_3 are 500 Hz, 1500 Hz, and 2500 Hz, respectively, for a vocal tract length of 17 cm. All other vowels are regarded as perturbations in different directions from the "neutral" schwa. An algorithm has been defined by

Öhman for calculating the formant frequencies of the continuous cross-sectional area curves of vowels that deviate from schwa. The formant frequencies corresponding to such deviation curves depend on whether the deviation curve is symmetric or antisymmetric about the midpoint. The symmetry of a curve is defined by taking the midpoint on the x-axis (distance from the glottis) and look at how the two parts of the curve to the left and right of the midpoint relate to each other. In a symmetric deviation curve the right part of the curve is a positive mirror image of the left part. If folded in half along a line midway from the glottis to the lips the right and left parts would cover each other. The deviation curve of [u] has a strong tendency to symmatry. In an antisymmetric curve the right half is negative mirror image of the left half. Both [i] and [a] have antisymmetric deviation curves.

Calculation of formant frequencies of a vocal tract tube with minimal termination impedance (i.e. unrounded lips) demonstrates two interesting facts. Firstly, any perturbation of such a vocal tract where the deviation curve is symmetric about the midpoint will have no acoustic effect on any of the formant frequencies. Secondly, the largest acoustic effects in such a vocal tract tube will be achieved by such perturbation of the vocal tract of schwa as result in deviation curves that are antisymmetric about the midpoint. In other words, moving the front of the tongue and the tongue-root in the same direction in relation to the roof of the mouth and the back pharyngal wall will have no acoustic effect, as long as there is no liprounding. Given the same lip condition, large acoustic differences will be obtained by moving the tongue body and tongue-root in different directions in relation to the roof of the mouth and the back pharyngal wall. Notice at this point that in going from i to [a] (where [a] is taken to be a low vowel midway between cardinal [a] and [ɑ]), the articulatory configurations are mainly antisymmetric. So for conditions with closure at the glottis and no impedance at the lips the most efficient way to vary formant frequencies (i.e. vowel qualities) corresponds precisely to that of the so called front vowels.

Conditions with higher impedance at the lips (i.e. rounded or closed lips) have not been worked out in any detail. But it is known that if a tube is terminated with a constriction providing a considerable impedance, then symmetric perturbations do affect formant frequencies.
It is, however, worth pointing out that back vowels that are basically rounded, also have their main constriction around the midpoint of the vocal tract, so that their vocal tract configurations tend to be symmetric. Apparently, the most efficient way to vary vowel qualities with rounded lips is by means of symmetric configurations[2].

We are now in a position to discuss real vowel systems. The following is an attempt to explain some common patterns of vowel mergings that occur in Niger-Congo languages with tongue root harmony.

One common pattern in these languages is that the vowels /ị/ and /e/, and /ụ/ and /o/ have merged, so that an earlier nine vowel system has become a seven vowel system. It is interesting to note that these sound changes seem to start with the merging of the two front unrounded vowels. We find languages today where these two front vowels have merged, or are merging, but the back rounded ones are not. Akan constitutes an example of a language where /ị/ and /e/ are in the process of merging, but there is no sign of /ụ/ and /o/ merging. Figure 1 is a typical formant frequency chart of eight Akan vowels (/a/ is excluded). The vowels were pronounced in short utterances by one speaker, and each utterance repeated five times. There is complete overlap of the formant frequencies of /ị/ and /e/. The corresponding back vowels /ụ/ and /o/ are kept acoustically separate by the second formant. It is worth noticing here that in Akan /ị/ and /e/ do not seem to contrast in stems, while /ụ/ and /o/ do.

Figure 2 shows typical tongue shapes of /ị/ and /e/ superimposed on each other, and of /ụ/ and /o/ superimposed on each other. Note that the highest point of the tongue in the vowels [ị] and [ụ] which are traditionally called high vowels, is much the same as it is in the vowels [e] and [o], which are traditionally called mid vowels. Although the tongue-shapes of /ị/ and /e/, and of /ụ/ and /o/ differ in very much the same way, the corresponding acoustic effects do not differ in the same way. The lack of a one-to-one correspondence between articulatory configurations and acoustic results is of course well known, and the more interesting question is the specification of what kinds of tongue shapes result in the same formant frequencies, and what kinds result in different formant frequencies. The phonologically mid /e/ and /o/ have an advanced tongue-root, and thus a larger pharyngal cavity than the phonologically high /ị/ and /ụ/. The mid /e/ and /o/ also have a lower front of the tongue, and thus a larger mouth cavity than the high /ị/ and /ụ/. Both halves of the vocal tract tube are larger in the mid vowels than in the high vowels. In other words, both /ị/ and /e/, and /ụ/ and /o/ are produced by perturbing the neutral schwa in such a way that the deviation curves are symmetric about the midpoint. This fact now provides an explanation as to why /ị/ and /e/ should merge, but not /ụ/ and /o/, as perturbation theory predicts that symmetric deviation curves will not differ as to formant frequencies for unrounded lip conditions, but will for rounded vowels. The deviation curves of all four vowels are symmetric, but only /ị/ and /e/ are unrounded. Therefore /ị/ and /e/ predictably have the same formant frequencies, while those of /ụ/ and /o/ differ because of the lip-rounding. Naturally, vocal tract shapes from real speakers will not be perfectly symmetric about the midpoint. In this case the high and mid vowels are approximately symmetric, and the acoustic effect is overlapping of the formant frequencies for the unrounded vowels, and non-overlapping for the rounded ones.

By relating the independently developed perturbation theory to phonetics the merging of /ị/ and /e/ has now been provided with an explanation that does not require the use of adhoc concepts like "marked - unmarked", and the like. I propose that in all those Niger-Congo languages where /ị/ and /e/ have merged, or are in the process of merging, the reasons are not to be found in looking for the tongue shapes becoming the same, but the explanation for the merging has an acoustic basis. Because of their both having symmetric tongue shapes they become acoustically the same, and are therefore starting to be perceived as the "same" vowel.

So why do earlier nine-vowel systems mostly become seven vowel systems by also merging /ụ/ and /o/ later, and not just eight vowel systems? I have no neat explanation for this, but I suggest that once /ị/ and /e/ overlap acoustically, this will create a structural pressure towards making the systems symmetric again by merging the corresponding back vowels. It is a fact that Niger-Congo vowel systems with harmony have a strong tendency towards symmetry. To sum up: a common development from a nine vowel system to a seven vowel system starts by unconditional merging of /ị/ and /e/ for acoustic reasons, then /ụ/ and /o/ merge for reasons of structural pressure towards symmetry.

When the tongue-root mechanism is involved in vowel production there is of course a very good possibility that /ị/ and /e/ will be articulated with symmetric tongue shapes and therefore merge, as has happened in many Niger-Congo languages. But there is another possibility. The vowels /ị/ and /e/ will merge, unless the pressures of communication within the language act to prevent this from happening. If a speaker wants to keep these two vowels distinct, he can easily do so by changing the tongue shape of /e/ to a more assymmetric shape. This can be accomplished by the following strategies: just lowering the front of the tongue, or just retracting the tongue root, or combining the two gestures. In the first and third case the acoustic effect is an increase of F_1. Published sources do not show how a pharyngal decrease by itself would affect the formant frequencies, but it seems likely that it would have the same effect, namely an increase of F_1. So whatever adjustment towards an asymmetric tongue shape a speaker chooses, the effect will be vowel lowering, and /e/ becoming more [ɛ]-like in quality. It might even merge with /ẹ/.

As there is no problem in keeping the rounded vowels with symmetric shapes separate, one would expect /u/, /ụ/, /o/ and /ọ/ to stay intact.

The pattern discussed above, where /ị/ and /e/ are kept separate, but /e/ has merged, or is merging, with /ẹ/, is exhibited by the development from Proto-Lower Niger to modern Lower Niger (Williamson, 1975). Most dialects of Igbo have an eight vowel system that is considered to be curiously skewed:

i   u
   ị   ụ
e[ɛ]   o
   a   ọ

K. Williamson (1975) posits a ten vowel system in Proto-Lower Niger. The Onitsha dialect of Igbo has the same phonological eight vowel system as most dialects of Igbo, but it has phonetically a nine-vowel system:

i   u
   ị   ụ
e   o
   [ɛ] ọ
      a

e and [ɛ] do not contrast: /e/ is realized as [ɛ] before tongueretracted vowels, as [e] elsewhere. Most dialects of Igbo and the Onitsha dialect can be interpreted in terms of the above discussion. If Igbo /ị/ and /e/ are to be kept separate, the tongue shape of /e/ must become more asymmetric. The distribution of [e] and [ɛ] in Onitsha indicates that these speakers have chosen to retract the tongue root of /e/ as the particular strategy for making /e/ more different from /ị/. In another dialect of Igbo, namely the Umuchu dialect, where radiographic evidence is available (Lindau, forthcoming), it is clear that the speaker has lowered /e/ by lowering the body of the tongue but retaining the tongue root distinction. (See figure 3.) At an earlier stage in the development from Proto-Lower Niger to today's Lower Niger languages the tongue-root advanced /a/ merged with /ạ/ (as in Onitsha) or with /ẹ/ (as in Ịka). As expected from perturbation theory these unconditioned mergers have affected only unrounded vowels, while the rounded vowels remain unaffected.

Looking at the Igbo vowel system in the light of perturbation theory thus explains the apparently "unnatural" skewness as a quite natural system, arising from an original "desire" to keep /ị/ and /e/ distinct, without making /e/ "higher" than /ị/.

Notes

1. The following description of perturbation theory is summarized from Öhman (1973).

2. There is as far as I know no evidence of what happens with rounded lips and antisymmetric tongueshapes. It is worth noting however, that front rounded vowels do differ acoustically, but the acoustic space of the front rounded vowels is considerably smaller than that of the front unrounded.

References

Fant G. 1960. Acoustic Theory of Speech Production. The Hague, Mouton

Lindau M. 1973. "Different vocal tract shapes with similar formant frequencies". Papers given to the 86th meeting of the Acoustical Society of America

Lindau M. Forthcoming. UCLA Ph. D. Dissertation

Öhman S. 1973. Untitled ms. UCLA

Stewart J.M. 1971. "Niger-Congo: Kwa". In Sebeok (ed.), Current Trends in Linguistics, Vol. VII: 179-212

Williamson K. 1975. "Proto-Lower Niger vowels: some first thoughts". Seminar paper, Department of Linguistics and Nigerian Languages, University of Ibadan

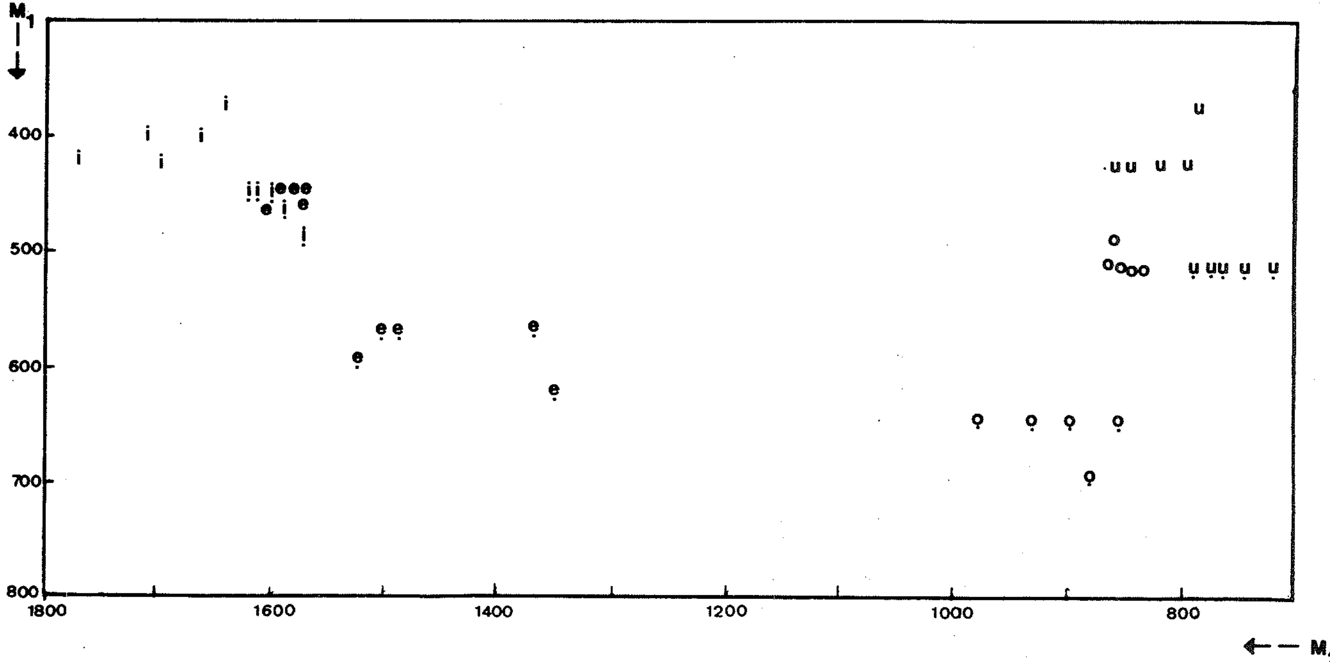


Figure 1. Formant chart with pitches in mel of eight Akan vowels of one speaker.

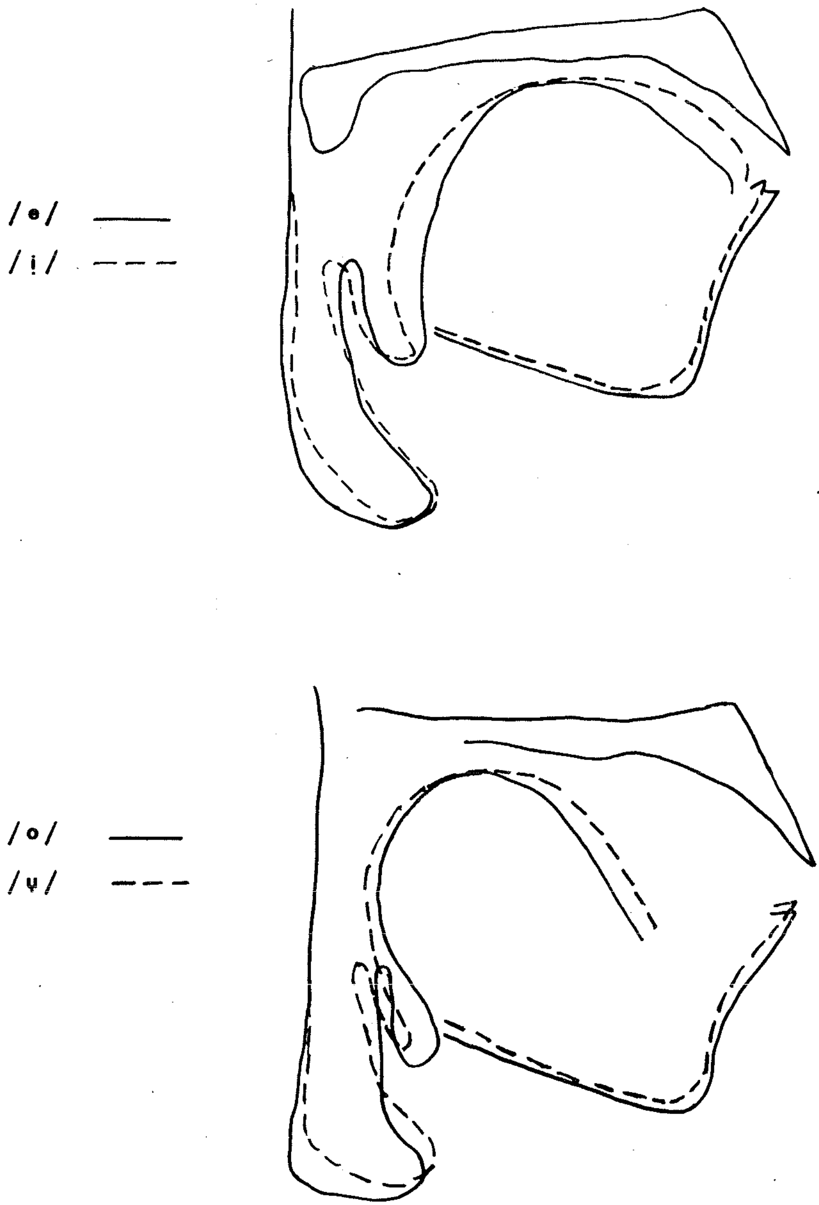


Figure 2. Tongue shapes of /e/ and /ị/, and /o/ and /ụ/ of a spekaer of Akan.

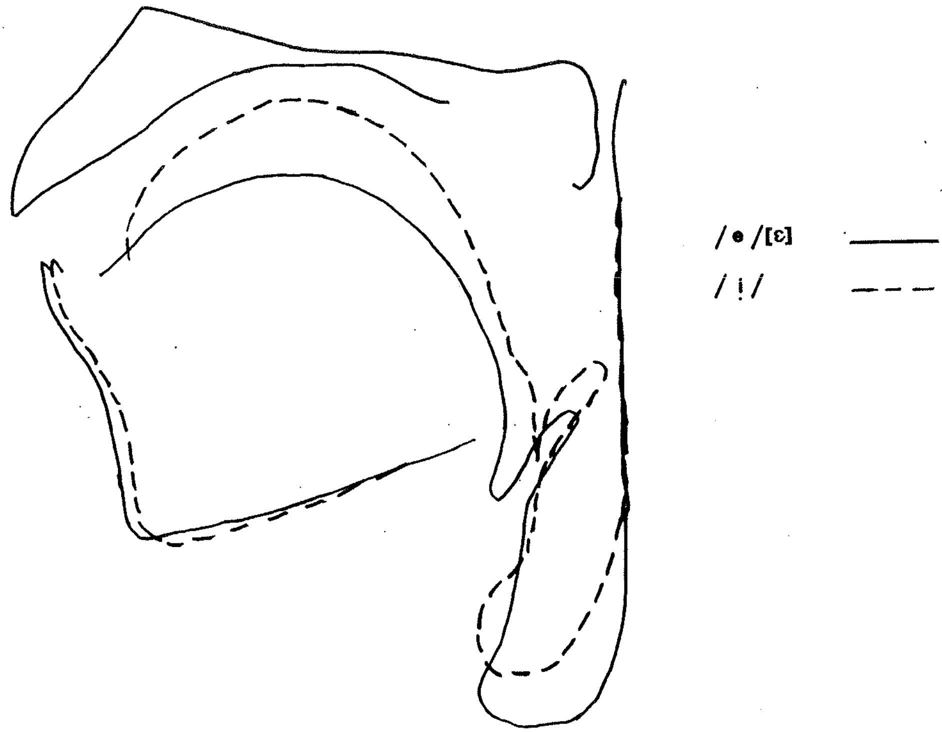


Figure 3. Tongue shapes of /e/ ([ɛ]) and /i̩/ of a speaker of the Umuchu dialect of Igbo.