Reprinted from

PMLA

PVBLICATIONS OF
THE·MODERN·LANGVAGE·ASSOCIATION·OF·AMERICA

Issued Six Times A Year

VOLUME LXVI SEPTEMBER, 1951 NUMBER 5

THE PHYSIOLOGICAL INTERPRETATION OF SOUND SPECTROGRAMS

PIERRE DELATTRE

Published by the MODERN LANGUAGE ASSOCIATION OF AMERICA
100 Washington Square East
New York

THE PHYSIOLOGICAL INTERPRETATION OF SOUND SPECTROGRAMS

By Pierre Delattre

In the physiological study of speech articulations our most objective information has come, until recently, from radiograms.[1] Now we have, in addition, spectrograms,[2] which, if we learn to interpret them, can also give us very objective information. For the typical phase (portion) of a speech sound, the interest of a spectrogram may be about equal to that of a radiogram; but for the transitional phases, the interest of a spectrogram will probably be much superior; and from the practical viewpoint of availability, the spectrogram will have a marked advantage for it can be had in a few minutes and at low cost. But to the linguist, the usefulness of a spectrogram depends on his ability to interpret it in articulatory terms. We need not stress, therefore, the importance of investigating the relation between formant[3] positions and speech organ positions at this stage of the still young science of sound spectrography.

Since Martin Joos's presentation of the relation between the articulatory triangle and the acoustic triangle in terms of a relation between formant 1 and tongue-height and between formant 2 and back-to-front tongue placement,[4] some progress must have been made by spectrography researchers. This is our contribution to this progress.

It is generally believed that "the shape of the filtering cavity is so very complex as to be mathematically unmanageable . . . "[5] This may seem very discouraging, but it is not. To the phonetician it means only that it will probably not be possible to determine exactly to what extent a certain formant can be assigned to a certain cavity. This is not of great importance to him. What he needs to know is rather the relation of each formant to the position or movement of articulatory organs. And that can be determined to a large extent (a) by studying the effect of *isolated* articulatory movements on formant positions with the help of spectrograms, (b) by comparing the spectrograms with radiograms, and in certain cases (c) by checking the findings on a speech synthesizer to see if the formant changes (or change) resulting from an isolated articulatory movement produced the auditory impression expected from this movement.

We shall apply this (a), (b), (c) technique successively to formants 1, 2, and 3 in that order. Before starting, let it be well understood that we are not concerned here with mouth cavities; to formant frequencies, we are exclusively trying to relate articulatory movements and positions. (More will be said about this in the discussion of formant 2.)

[1] X-ray pictures of the organs of speech during articulation.

[2] Spectrographic pictures of the acoustic resonances produced by the speech organs during articulation, in three acoustic dimensions: time, frequency, intensity. The first extensive presentation of such pictures is to be found in *Visible Speech* by Potter, Copp, and Green (New York: Van Nostrand, 1947). Briefly, a sound spectrogram shows the energy distribution on a time-frequency scale where time is read from left to right, frequency from bottom to top, energy by the degree of darkness.

[3] For those who are not yet familiar with spectrography, we shall define the essential term *formant* as it is used here. Linguistically the color of a vowel is determined by the frequency position of its formants—mainly its two lowest formants. Let us look at Fig. 1 or Fig. 3. There, formants appear as dark horizontal bands on a linear frequency scale (range: 3500 cycles from bottom to top). For instance, for [œ] ,the lowest band is formant 1 (frequency: about 500 cycles), the one above is formant 2 (frequency: about 1400 cycles), and the next one above is formant 3 (frequency: about 2400 cycles). Thus, on our spectrograms formants appear as the darkest areas. Acoustically, formants are the frequency regions of greatest intensity. For voiced vowels, the number of harmonics that cross such regions (in other words, that are comprised in formants) usually vary from one for high female voice to two or three for male voice. The frequency of a formant can satisfactorily be given by the frequency of its center.

[4] *Acoustic Phonetics*, Linguistic Society of America (Baltimore, 1948), pp. 49–59.

[5] Ibid., p. 57.

Formant 1

The phonetic triangles and quadrilaterals, in their vertical direction, have all been based on tongue-height (the highest point of the tongue arch); therefore, it was natural for Joos to speak of tongue-height when relating the triangle obtained by plotting formants 1 and 2 on a logarithmic scale to the traditional phonetic (articulatory) triangle. However, the term tongue-height may not be the most appropriate in relation to the frequency of formant 1. Perhaps the more general term of "opening" (meaning overall opening of the oral tract with definite relation to the width of the strictures at the main points of articulation, but not depending entirely on that) would be more correct. If we examine formant 1 not only on spectrums of vowels but also on those of voiced consonants, this will appear clearly.

Vowels. Let us compare the two vowels [i] and [u], for instance. According to our measurements their formant 1 frequencies are nearly the same. In many cases they are exactly the same, as in the series of spectrograms on which our French acoustic triangle is based. See Fig. 1: [i] and [u] have same formant 1 frequencies on the spectrograms (bottom left); consequently the line joining them on the acoustic chart (center) is horizontal. It does not sink to the right as in the Jones quadrilateral (Fig. 2). The same can be seen on Fig. 4 where formant 1 follows a straight horizontal line from [i] to [u]. Confirmation of this frequency similarity for [i] and [u] formant 1 can also be obtained by the synthetic production of those vowels. On the Cooper pattern playback,[6] with the harmonic channels set at 120 cycles apart, the best [i] and the best [u] are both produced when formant 1 is centered around the second harmonic, at 240 cycles. If without changing formant 2 for [u], formant 1

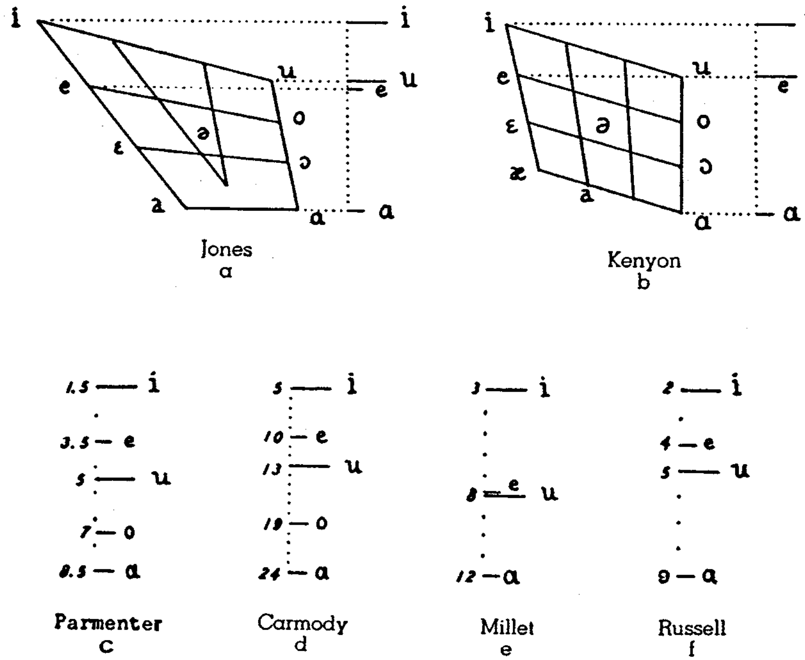

FIG. 2. Tongue-height comparisons for [i] and [u] according to six different sources.

is raised to the third harmonic (360 cycles)—which is the best harmonic for [o]—the result is a sound much closer to [o] than to [u]. In fact, the contrast between [u] and [o] seems to depend more on changing the position of formant 1 than that of formant 2.

This similarity of [i] and [u] formant 1 frequencies does not correspond

[6] The *Pattern Playback*, developed by Franklin S. Cooper at the Haskins Laboratories, New York, is a speech-synthesizer that permits us to transform hand-drawn spectrograms into sound, using modulated light that is reflected from hand-drawn white lines. The relative intensity of each harmonic or of each formant depends on the width of the lines drawn in the harmonic channels. For our use of the machine, the harmonic channels were set 120 cycles apart and there were 50 channels for a total frequency range of 6000 cycles. Among its many uses, this machine makes it possible to study the effects upon speech obtained by omitting or adding some resonances, or by modifying either their intensity, their frequency or their type.

to *tongue-height* as known through radiography. The articulatory quadrilaterals of Jones[7] and Kenyon[8] both give [u] a tongue-height much lower than that of [i] and very nearly as low as that of [e] (Fig. 2, a and b). And this tongue-height difference between [i] and [u] is not even as marked as in actual sets of x-rays that we have consulted. Parmenter's radiograms of American vowels[9] show totally different tongue-heights for [i] and [u]: if we measure the distance in mm. from the highest point of the tongue to the highest point of the palate, we find that, with [a] at 8.5 and [e] at 3.5, [i] is at 1.5 and [u] at 5.0 (Fig. 2c). On the Holbrook-Carmody x-rays of superimposed vowels from many languages,[10] the same measurements (on another scale) give: with [a] at 24 and [e] at 10, [i] at 5 and [u] at 13 (Fig. 2d). On the Czech vowels of M. Hala, as presented by Adrien Millet,[11] similar measurements on a different scale give: with [a] at 12 and [e] at 8, [i] at 3 and [u] at 8 (Fig. 2e). On the French vowels x-rayed by Oscar G. Russell,[12] with [a] at 9 and [e] at 4, [i] is at 2 and [u] at 5 (Fig. 2f). All these radiograms concord rather closely in presenting a considerable difference in tongue-height between [i] and [u]—even more difference than on the Jones quadrilateral.

If we compare two other vowels, such as [e] and [o], in the same manner, we find almost the same aspect. Briefly, [e] and [o] have very nearly the same frequencies for formant 1 on the spectrograms, and their synthetic production is done satisfactorily by using the same formant 1 frequency for both; but their tongue-heights show differences about comparable to those between [i] and [u]. For instance, on the Parmenter x-rays, with [a] at 8.5, [e] is at 3.5 and [o] at 7.0. On the Holbrook-Carmody x-rays, with [a] at 24, [e] is at 10 and [o] at 19.

All this confirms that tongue-height differences and frequency differences of formants 1 for different vowels do not correspond when comparing articulatory triangles with acoustic triangles. If we look for some feature that is nearly the same for [i] and [u], for [e] and [o], in the articulation, we find it better in the overall opening of the mouth tract, as measured for instance by the distances between the upper and lower incisors or by the distances between the highest point of the tongue and the point of the palate closest to it—in other words, the general width of the structures. Either of these two measurements show only small

[7] Daniel Jones, *An Outline of English Phonetics* (Cambridge: Heffer, 1936), p. 63.

[8] John S. Kenyon, *American Pronunciation* (Ann Arbor: Wahr, 1937), p. 66.

[9] C. E. Parmenter and C. A. Bevans, "Analysis of Speech Radiographs," *American Speech*, VIII, iii, 51.

[10] Richard T. Holbrook and Francis J. Carmody, "X-ray Studies of Speech Articulations," *Univ. of Calif. Publ. in Mod. Philol.* XX, iv, 230.

[11] *L'articulation des voyelles* (Paris; Vrin, 1937), p. 7.

[12] *The Vowel* (Columbus: Ohio State Univ. Press, 1928), pp. 110–111.

differences between corresponding front and back vowels. The use of the term *opening* instead of *tongue-height* seems indicated.

CONSONANTS. The voice bar—or formant 1—of voiced consonants serves to further confirm that tongue-height is not the most appropriate term in relation to formant 1 frequency. First we notice, by comparisons with one or another of the consonants that have the most clearly delimited voice bars, that the frequency of formant 1 is in accord with the accepted notion of degrees in opening of the mouth: judging from a large set of consonant spectrograms we made at the Bell Telephone Laboratories of New York (summer of 1947) (Fig. 3), the frequency of formant 1 for [b] and [d], is always lower than for [m] and [n], and the latter is always lower than for [l].

Secondly, we notice that the frequencies of formant 1 are the same for [b] as for [d] ,whereas for [b], a bi-labial, the tongue position is not involved as it is in [d]. Similarly, the frequency of formant 1 is the same for [m] as for [n], whereas the tongue is not involved in [m] as it is in [n]. What [b] and [d] have in common, what [m] and [n] have in common, is not a certain degree of tongue-height but a certain degree of overall closure.

Conclusion for formant 1. The relation between formant 1 position and articulatory position should be stated in the following terms: There is a direct relation between formant 1 frequency rising and overall opening of the oral tract. The higher the frequency of formant 1, the wider the overall opening; and inversely.

Formant 2

Two introductory remarks will prepare for the discussions that follow.

A. One weakness of the traditional phonetic triangle (or of the Jones quadrilateral) is that it is based exclusively on tongue positions. It leaves out any information about color difference that is due to other causes such as lip spreading-to-rounding. Thus it usually places the front rounded vowels on top of the front spread ones, and the back rounded vowels on top of the back spread ones, because it assigns them respectively the same tongue positions. The *acoustic* triangle (or quadrilateral) does not leave out lip rounding: it gives a different place to rounded vowels than to spread vowels that have the same tongue placement (see Fig. 1).

B. A second weakness of the traditional physiological triangle is that it measures tongue backing by the highest point of the tongue on X-ray profiles. This highest point does not necessarily agree with back-and-up profiles. The *acoustic* triangle is probably based on actual retraction of the tongue. The *acoustic* triangle is probably based on actual back-and-up tongue retraction, whether or not it agrees with the highest point of the tongue.

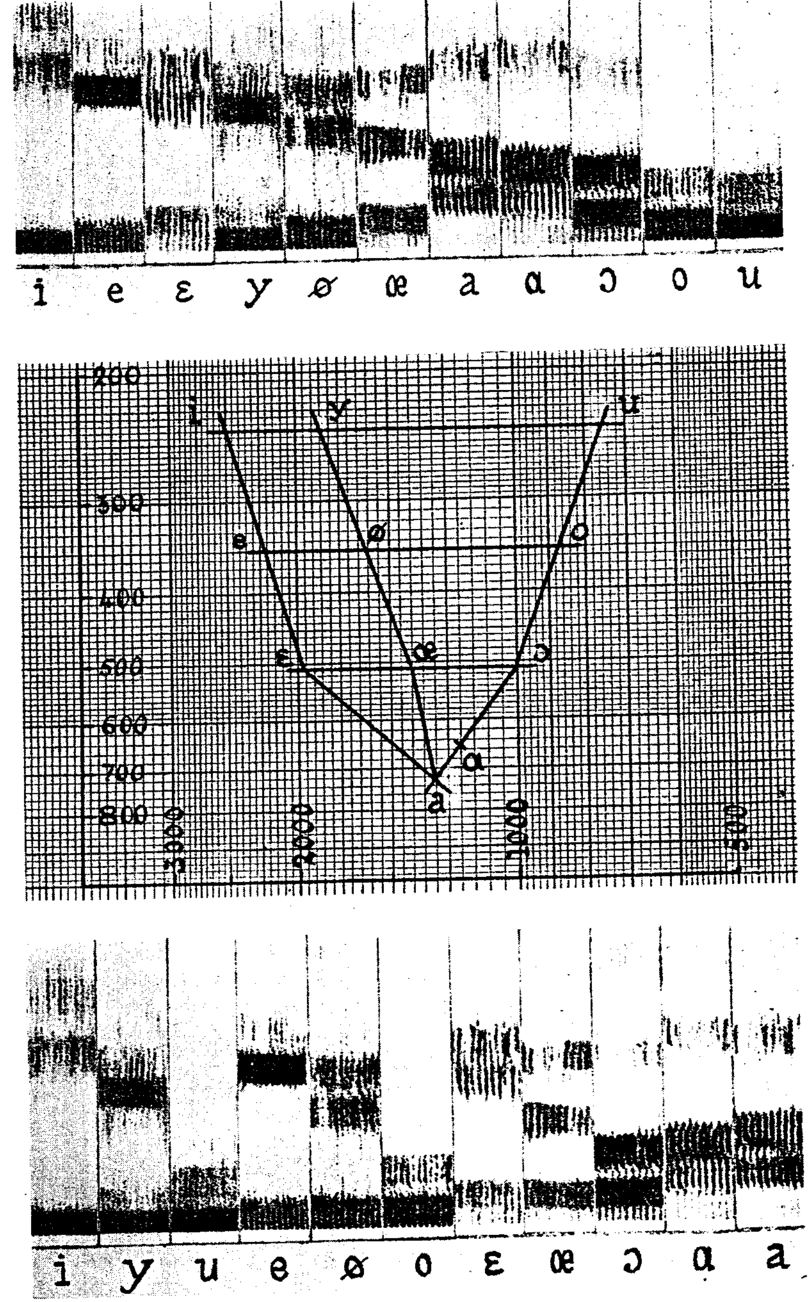

FIG. 1: AN ACOUSTIC VOWEL CHART. This chart is reproduced from my article, "Un triangle acoustique des voyelles orales du français," *French Review*, XXI (May 1948), 481. The place of each vowel on the chart was determined by plotting the frequency of formant 1 vertically versus the frequency of formant 2 horizontally. Plotting is done on a logarithmic scale in order that the relative distances from one vowel to another correspond to the auditory impression and not to the acoustic frequency. This way, equal intervals on the chart correspond to equal intervals for the ear. Below the triangle, spectrograms are arranged to show the order in which the frequency of formant 1 increases (oral tract opening). Above the triangle, they are arranged to show the order in which the frequency of formant 2 decreases (front cavity lengthening).

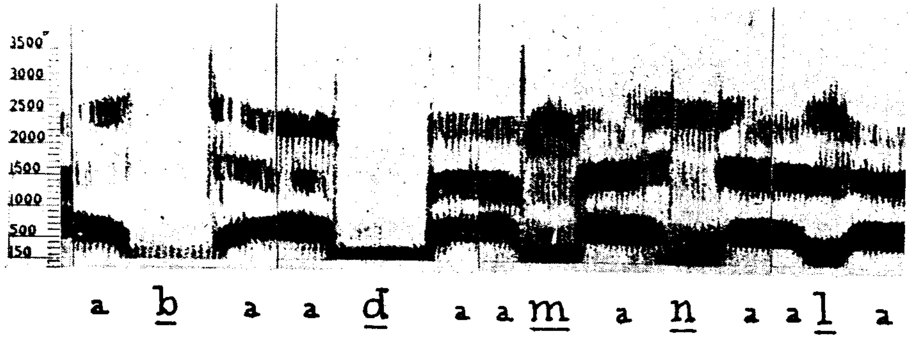

FIG. 3. Spectrograms of [b], [d], [m], [n], [l], showing the frequency of formant 1 rising from left to right in three steps: [b d], [m n], [l]. (Scale is disposed for reading measurements at center of formants.)

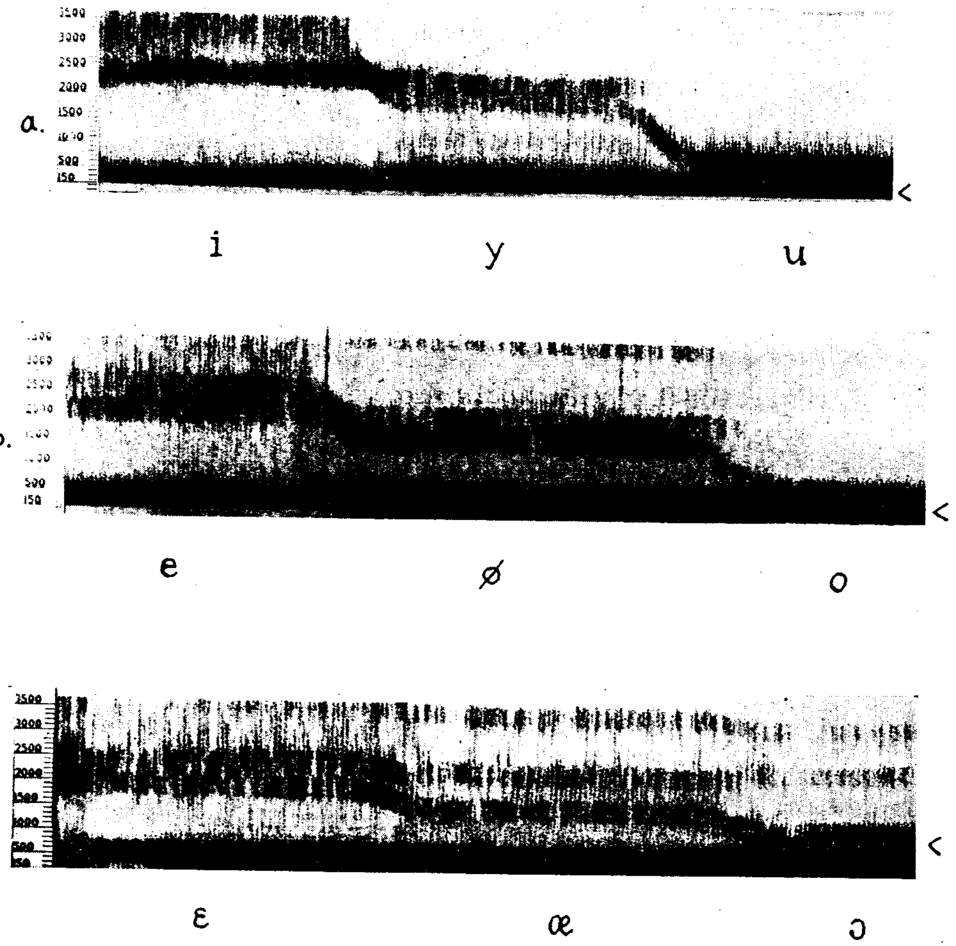

FIG. 4. Spectrograms showing the lowering of formant 2 frequencies, either by lip rounding: [i]–[y], [e]–[ø], [ɛ]–[œ]; or by tongue backing: [y]–[u], [ø]–[o], [œ]–[ɔ]. (Scale is disposed for reading measurements at center of formants.)

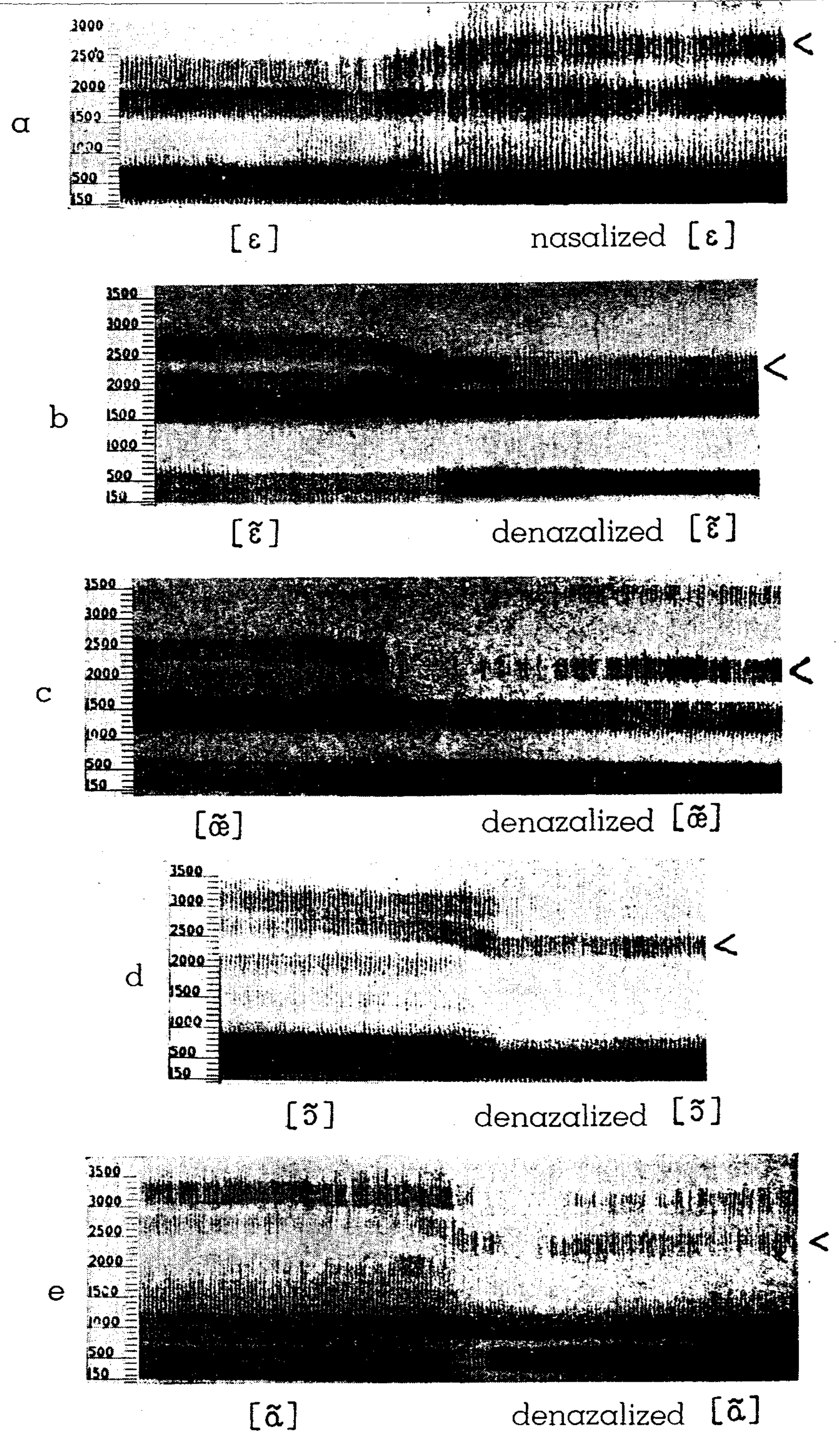

FIG. 6. Spectrograms showing the 300 cycle shift of formant 3 when the velum is lowered (from left to right) as in *a*, or raised as in *b*, *c*, *d*, *e*, all other organs being kept as immobile as possible. (Scale is disposed for reading measurements at center of formants.)

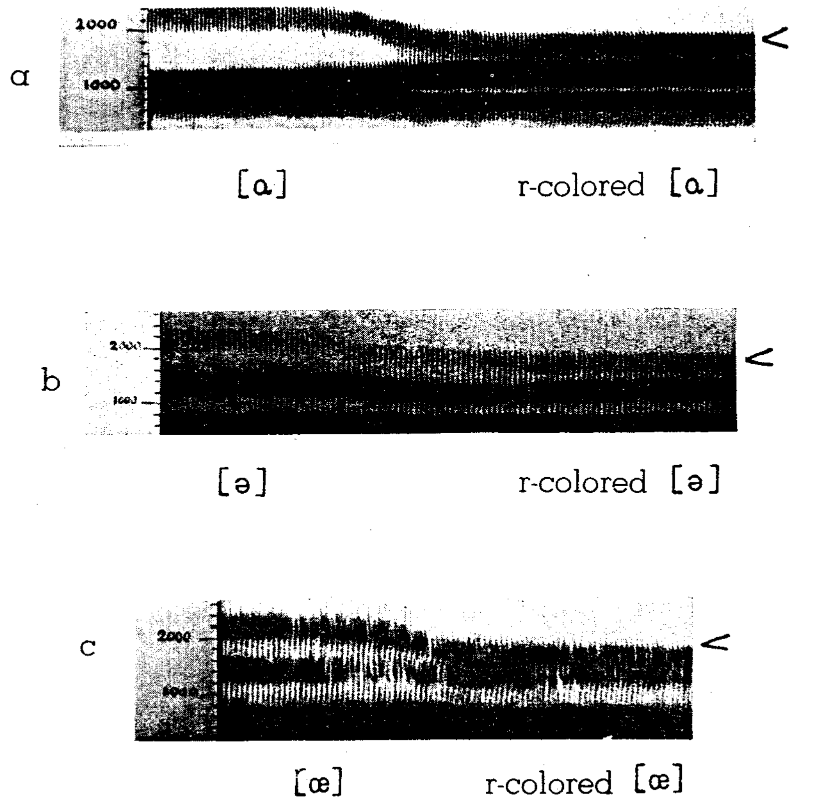

FIG. 7. Spectrograms showing the shift of formant 3 when the tongue tip is raised (from left to right) as for Midwestern American *r*. (Scale is disposed for reading measurements at center of formants.)

In those two remarks lie the main causes of discrepancy between the traditional vowel charts (articulatory) and the acoustic vowel charts such as that of Fig. 1.

Joos, with the traditional tongue placement triangle in mind, has already made clear the general relation that exists between formant 2 and back-to-front tongue placement. We wish 1) to further discuss this relation of formant 2 with back-to-front tongue placement; 2) to establish another striking relation of formant 2, the one it has with lip spreading-to-rounding; 3) to discuss the conjugating of these two notions into a single one.

1. We shall see here that there is a direct relation between tongue backing and formant 2 frequency lowering, but not quite in the same sense as is implied by the front-to-back horizontal direction of the phonetic triangle or quadrilateral. In relation to formant 2 lowering, tongue backing is not measured by how far back the highest point of the tongue is (as in the phonetic triangle) but by how far back-and-up the tongue as a whole is retracted. This retraction cannot be measured as well on radiograms as the highest point of the tongue but it does show clearly (as we shall demonstrate farther on), and besides it is felt kinesthetically much better than the highest point of the tongue (one feels that the tongue is pulled back-and-up more for [u] than for [ɔ]).

To exemplify this we should compare pairs of vowels that differ by tongue backing only, all other conditions remaining practically equal. Such are the three French pairs: [y]-[u], [ø]-[o], [œ]-[ɔ] (Fig. 4). From [y] to [u] (Fig. 4a), the lips remain equally rounded and the jaws equally closed; the only important change is in the tongue position which passes from extreme front-and-up to extreme back-and-up. On the spectrogram this backing of the tongue is translated by a marked lowering of formant 2 while formant 1 remains in the same place. Exactly the same could be said of the transitions from [ø] to [o] (Fig. 4b) and from [œ] to [ɔ] (Fig. 4c).

Comparisons can also be made satisfactorily regardless of the degree of opening as long as the vowels agree from the angle of rounding-spreading. In the series [i], [e], [ɛ] (Fig. 1, top left), where all the vowels have lip spreading, formant 2 lowering is in accord with tongue backing. In the French series [y], [ø], [œ] (Fig. 1, top center) and [ɔ], [o], [u] (Fig. 1 top right), where all the vowels have definite lip rounding, formant 2 lowering again goes with tongue backing.

For the last three vowels [ɔ], [o], [u], however, the traditional vowel quadrilateral does not agree with the acoustic feature of formant 2 lowering (compare [ɔ], [o], [u] on Fig. 2 with [ɔ], [o], [u] on Fig. 1). The Jones quadrilateral shows [u] less far back than [o], and [o] less far back than [ɔ] (the way the charts are generally disposed, [u] is left of [o], and [o] left of [ɔ]). There are two striking reasons for this discrepancy between the Jones chart (Fig. 2) and the acoustic chart (Fig. 1). (a) The Jones chart does not take into account the important feature of rounding, which, as we shall see next, also has a lowering effect on formant 2. For instance, if [u], [o], [ɔ] had the same tongue backing (as is the case on simplified quadrilaterals), on the acoustic chart, [o] would still be on the right of [ɔ] because it is more rounded, and [u] would still be on the right of [o] for the same reason. (b) The Jones chart bases its notion of tongue backing on the backing of the highest point of the tongue arch and not on actual back-and-up retraction of the tongue as it is felt kinesthetically.

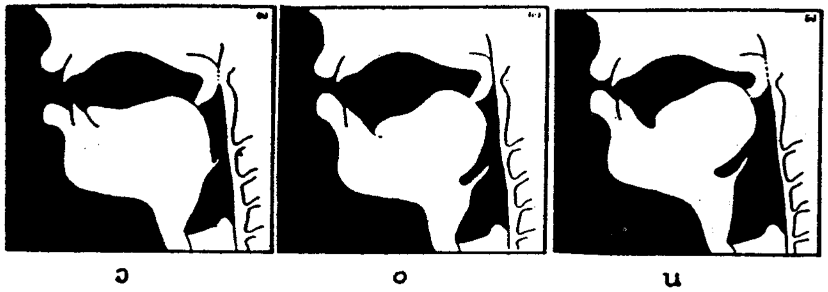

FIG. 5. Radiographs showing the back-and-up retraction of the tongue from [ɔ] to [u]. These radiographs are reproduced with the permission of the authors, from an article entitled "Analysis of Speech Radiographs," by C. E. Parmenter and C. A. Bevans, *American Speech*, VIII, 3, p. 51.

In order to verify this last statement let us examine some radiograms of [ɔ], [o], [u] presented by Parmenter and Bevans (Fig. 5). The 3 pictures show very nearly the same distance from the teeth to the highest point of the tongue, yet how different is the mass of the tongue! Look at the tongue tip, especially. For [ɔ] it lies flat almost reaching the top of the lower incisors. For [o] it points toward the roots of the incisors. For [u] it disappears into the mass of the tongue. Comparing [ɔ] to [u], we are bound to feel kinesthetically the difference in back-and-up tongue retraction that is so eloquently shown by these X-rays. Therefore, we may conclude that there is a direct relation between formant 2 lowering and tongue backing if we estimate tongue backing not by the highest point of the tongue arch but by the back-and-up retraction of the tongue as it is felt kinesthetically.

2. Let us examine the relation between lip rounding and formant 2 lowering. To study it we may take any pair of vowels in which the two sounds differ by lip rounding only and are about similar otherwise. Fig.

4 offers us 3 such pairs. [i] and [y] have about same opening and same tongue fronting, but [i] has spread lips and [y] rounded lips: passing from [i] to [y], a clear lowering of formant 2 can be observed on the spectrogram (Fig. 4a). The same can be said of the pairs: [e]-[ø] (Fig. 4b), and [ɛ]-[œ] (Fig. 4c).

We may conclude that there is a direct relation between lip rounding and frequency lowering of formant 2: as the lips are rounding formant 2 is lowering, and inversely.

3. Trying to find a common denominator for the two preceding relations (tongue backing and lip rounding) to formant 2 lowering, we noticed that both tongue backing and lip rounding had the effect of lengthening the front (or mouth) cavity (Fig. 4a, b, c): the *highest* frequency for formant 2 is obtained for vowel [i] which seems to have the shortest possible front cavity (with maximum tongue fronting and maximum lip spreading); the *lowest* frequency for formant 2 is obtained for vowel [u] which seems to have the longest possible front cavity (with maximum tongue backing and maximum lip spreading). We stated this at the MLA meeting of 1947 (in our paper on the nasal resonances of French nasal vowels) and again in an article that was to serve as a simple introduction to spectrography for French teachers. "Il existe une relation constante et inverse entre la hauteur de la formante 2 et la longeur de la cavité de résonance buccale."[13] To put this in terms similar to those that we have been using here we should say: there exists a direct relation between formant 2 lowering and front cavity lengthening. This is a good statement only if "cavity lengthening" is interpreted appropriately. Cavity lengthening is not an acoustic feature here but a physiological one: using the terminology "cavity lengthening" permits to include two physiological movements (tongue backing and lip rounding) into one statement. In fact it must mean tongue backing and/or lip rounding. It does not imply any notion as to the size of the cavity. For instance, should it imply that by lengthening the cavity becomes larger it would also imply that formant 2 is independently related to the size of the front cavity. And that is not true or at least is not known to be true: because the speech cavities are "mathematically unmanageable"[5] we don't know to what extent formant 2 is related to front cavity; we only know that theoretically any formant is related to the whole system of speech cavities. It is probable that the relation between formant 2 and front cavity varies for each vowel. The opinion of H. K. Dunn of the Bell Telephone Laboratories is that the farther apart formants 1 and 2 are the more they can be assigned respectively to the back and the front cavities, and the nearer

[13] Pierre Delattre, "Un triangle acoustique des voyelles orales du français," *French Review*, XXI (May 1948), 477–484.

they are to one another the more they both must be assigned equally to both cavities. This is about the limit of what can be known at present on the degree of independent relations between cavities and formants.

Let us summarize our findings for formant 2.

1. There is a direct relation between back-and-up tongue retracting and formant 2 frequency lowering: the more the tongue is retracted the more the frequency of formant 2 is lowered: and inversely.

2. There is a direct relation between lip rounding and formant 2 frequency lowering: the more the lips are rounded (and protruded) the more the frequency of formant 2 is lowered; and inversely.

3. Since tongue backing and lip rounding (a) both tend to lengthen the front cavity of the mouth and (b) both have a lowering effect on formant 2 the relations expressed in the two preceding paragraphs can be conjugated into one, to say: There is a direct relation between front cavity lengthening and formant 2 frequency lowering: the longer the front cavity the lower the frequency of formant 2; and inversely. "Cavity lengthening" is not an acoustic feature, here, but a physiological one; it has two main factors: tongue backing and lip rounding. (The mathematics of resonant cavities will show that lip rounding ought to have this effect: narrowing of a cavity opening will lower its resonant frequency and thus counterfeit lengthening.)

Formant 3

First a few words of introduction since formant 3 is not so well known as the two main formants 1 and 2. Synthetic speech experiments furnish us much of the following information.[14] No doubt that formant 3 is much less responsible than formants 1 and 2 for the linguistic color of vowels. Formant 3 is mainly to be considered as one of the many higher resonances that show on the spectrum of any vowel. Being the lowest of these, it has the most perceptible effect. But that is not saying much; for as a whole these high resonances above formant 2 have very little effect on color, they mostly add intelligibility without changing the color and are probably responsible for voice quality. The contribution of formant 3 to the color or intelligibility of vowels increases as formant 2 is higher. For a well fronted [i], it appears to be as important as formant 2 in shaping the true color. For [o] and [u], its absence is hardly noticeable. For the others it affects the degree of intelligibility but hardly the distinctive color unless it is moved up or down from its normal position. If it is moved up slightly (as little as 200 cycles) it causes a small but perceptible change in color comparable to "more open," "more backed" (this is not the place to describe such changes) but not nasalization;

[14] See note 5.

if it is moved down, it adds r-color (midwestern r) to vowels. This last feature was mentioned by Joos[15] and we can confirm it after conclusive synthetic experiments: for instance, [a] reflects r-color when formant 3 is lowered below its normal level of about 2600 cycles, and r-color increases as formant 3 comes closer to formant 2 (cf. Fig. 7a).

Let us examine now two cases of simple relation between formant 3 and some *articulatory* movement.

1. Velum relation to formant 3. (The velum motions are involuntary. We do not feel them and cannot control them directly. But they obey our seeking to produce nasality and we can control them indirectly by doing just that. Therefore, "lowering the velum," here, will mean "seeking to nasalize.")

Starting from an oral vowel, if we lower the velum while holding all other speech organs immobile, the frequency of formant 3 rises considerably while the frequencies of formants 1 and 2 remain fairly stable. Fig. 6 shows this rise of formant 3 in nasalizing of an [ɛ].[16] (We are speaking only of frequencies, not of the intensities or modes of those formants, here. This rise of formant 3 averages very close to 300 cycles in the case of French nasal vowels as compared with oral ones. It is a little more marked for [ɑ̃] than for the three other nasal vowels [ɛ̃], [œ̃], [ɔ̃].

The opposite experiment confirms the one above. If, for instance, a French nasal vowel is denasalized by raising the velum while trying to hold all other organs immobile, the frequency of formant 3 lowers by some 300 cycles while the frequencies for formants 1 and 2 remain practically stable. Fig. 6b, c, d, e, show this lowering of formant 3 in the denasalizing of [ɛ̃], [œ̃], [ɔ̃] and [ɑ̃].[17]

Although this is apart from the subject, we must say here that the 300 cycle rise of formant 3 mentioned above does not seem to be related to the nasal quality of nasal vowels. If a nasal vowel is hand-drawn on the Cooper pattern playback and produced synthetically, whether formant 3 is drawn with its 300 cycle rise or without has no appreciable effect on the nasal quality of the vowel. Since the addition of the nasal quality to the vowel, when the velum is lowered can clearly be isolated in synthetic experimentation and can be assigned to other features than the 300 cycle rise of formant 3, we may be justified in saying that

[15] Op. cit., p. 93.

[16] The result, a nasalized [ɛ], is not to be confused with the real French nasal [ɛ̃], which does not have the same articulatory positions, hence the same formant 1 and 2 frequencies (apart from the velum lowering).

[17] The result of such denasalizing does not give French oral vowels [ɛ], [œ], [ɔ], [ɑ], but some strange vowels that do not exist in French (nor probably in any language), for the organic positions of the four French nasals (and their formants 1 and 2) are not the same as those of any French orals. This can be shown by synthetic speech as well as by human speech.

a) the formant 3 rise is an unavoidable effect of lowering the velum;

b) lowering the velum causes (by adding one more cavity to the others) several resonances and additions, some of which are clearly responsible for nasal quality;

c) the formant 3 rise is not one of the changes appreciably responsible for nasal quality; rather it has an effect on the *color* of the vowel, independently from its nasality, and comparable to the effect of formant 2. (An article on French nasal vowels soon to appear will treat this point in full.)

We conclude that there exists a direct relation between the frequency rising of formant 3 above its normal level and the lowering of the velum as it is lowered in nasalizing.

We limit our conclusion to the case where the velum is lowered "as in nasalizing," that is, with the back part of the velum away from the wall of the pharynx so as to allow the nasal cavity to communicate with the oral cavity. It is possible that lowering the velum without the extremity leaving the wall of the pharynx be also related to the rising of formant 3, but it is not probable. We experimented with French [ʀ], which requires the back of the tongue and the velum to draw toward one another. In the case of the front rounded vowels, the transitions to [ʀ] seemed to indicate, in addition to tongue backing shown by lowering formant 2, velum lowering shown by raising formant 3. So, it is very tempting to interpret the inverse sinuosities of formants 2 and 3 as the movements of tongue and velum drawing together. However this interpretation is not upheld by the examination of transitions of all other vowels to [ʀ]. For instance, the [a] transition to [ʀ] shows no rise of formant 3, and the [i] transition to [ʀ] even shows a lowering of formant 3.

2. We mentioned above how *r*-color can be added to [a] by lowering the frequency of formant 3 on a speech synthesizer such as the Cooper pattern playback. To add this *r*-color to any vowel, in human speech, it is sufficient to raise the tip of the tongue (let it take the well known retroflex position) while making no effort to change the positions of any other organs. Fig. 7 shows three transitions from an ordinary vowel to an *r*-colored vowel by simple raising of the tongue tip. The vowels of Figs. 7a and 7b were uttered by an American from Michigan. The vowel from Fig. 7c was uttered by a Frenchman. Spectrums of the tongue raising transition always show a frequency lowering of formant 3. The range of this frequency change is generally considerable; it may reach 1000 cycles.

Conclusion: There exists a direct relation between the frequency lowering of formant 3 below its normal level and the raising of the tongue tip toward a retroflex position as in the articulation of Midwestern-American *r*.

However the effect of tongue tip raising is not always limited to formant 3. Formant 2 is also affected in some cases. When formant 2 is already close to formant 3 as for French [y] and [ø], tongue tip raising lowers formant 2 alongside formant 3. More generally, when formant 2 is higher than for [ə], tongue raising tends to lower it toward its [ə] frequency, and when formant 2 is lower than for [ə], tongue tip raising tends to cause it to rise toward its [ə] frequency.

Formant 1 also seems to be affected by tongue tip raising but it really is not, at least not directly. When tongue tip is raised, formant 1 tends toward the frequency it has for [e], approximately. But this can be overcome by keeping the general opening of the mouth very stable. Therefore we may say that tongue tip raising affects formant 1 only if mouth closing or opening takes place at the same time.

Summary

We have tried to bring out the articulatory meaning of formants 1, 2, and 3 of sound spectrograms in a discussion of their relations with tongue-height, mouth opening, tongue backing, lip rounding, front cavity lengthening, velum lowering, and tongue tip raising; and to show that direct relations exist between the following formant frequency changes and the following articulatory movements: 1) between formant 1 frequency raising and overall mouth opening: 2) between formant 2 frequency lowering and tongue backing; 3) between formant 2 frequency lowering and lip rounding; 4) between formant 2 frequency lowering and front cavity lengthening, this being a manner of conjugating statements 2 and 3 into a single one (cavity length, here, has a physiological meaning only, not an acoustical one); 5) between formant 3 frequency raising and velum lowering as in nasalizing; 6) between formant 3 frequency lowering and tongue tip raising as in *r*-coloring.[18]

UNIVERSITY OF PENNSYLVANIA
Philadelphia 4

[18] This research was made possible by a University of Pennsylvania Faculty Research Grant to the author and by a Carnegie Corporation of New York Grant to the Haskins Laboratories, New York. We also wish to express our deepest appreciation to the staff members of the Haskins Laboratories and Bell Telephone Laboratories.