UCLA
Working Papers in Phonetics

Title
WPP, No. 31

Permalink
https://escholarship.org/uc/item/31f5j8m7

Publication Date
1976-03-01

eScholarship.org Powered by the California Digital Library
University of California

UCLA

1976

UCLA Working Papers in Phonetics 31

March, 1976

The UCLA Phonetics Laboratory Group

Research	Technical and Secretarial
Catherine Browman	
Ron Carlson	
Victoria A. Fromkin	Willie Martin
Linda Galloway	Anna Meyer
Jack Gandour	Renée Wellin
Louis Goldstein	
S. Robert Greenberg	
Richard Harshman	
Jean-Marie Hombert	
Dan Kahn	
Peter Ladefoged	
Mona Lindau	
Ian Maddieson	
George Papçun	
D. Lloyd Rice	
Georgette Silva	
Diana Van Lancker	

As on previous occasions, the material which is presented here is simply a record for our own use, a report as required by the funding agencies, and a preliminary account of work in progress.

Funds for the UCLA Phonetics Laboratory are provided through:

USPHS grant NS 9780
USPHS grant NS 12303
NSF grant SOC75-7158
Army contract DAAB0375C0371
and the UCLA Department of Linguistics

Correspondence concerning this series should be addressed to:

Professor Peter Ladefoged
Department of Linguistics
Phonetics Laboratory
UCLA
Los Angeles, California 90024

The phonetic specification of the languages of the world

Peter Ladefoged

[Paper presented at the Eighth International Congress
of Phonetic Sciences, 17 August 1975]

> *"I often say that when you can measure what you are speaking about and express it in numbers you know something about it; but when you cannot express it in numbers, your knowledge is of a meagre and unsatisfactory kind; it may be the beginning of knowledge, but you have scarcely in your thoughts advanced to the stage of science."*
>
> Lord Kelvin

The theme of this paper is Lord Kelvin's dictum (Thompson, 1889; Vol I, p. 73). In phonetics, as in other sciences, we do not really know anything until we can express it in terms of numbers. We might also add the other generally recognized scientific criteria: the numbers must be valid, reliable and significant. They must be valid in the sense that they are truly representative of the thing we think we are measuring. They must be reliable and not subject to instrumental inaccuracy or procedural artifacts. They must be statistically significant and not due to chance.

The problem for the phonetician is that there seem to be two completely independent ways in which descriptions of languages can be formulated in terms of numbers. Which system we use depends on our reason for making the description. If we are interested in characterizing each language as economically as possible, we will use one of the systems developed for speech synthesis. But if we are concerned with explaining why languages are the way that they are we will have to use one of the feature systems developed by linguists. To some extent the two procedures will produce sets of numbers that overlap, or at least can be predicted one from another. But for the most part the two procedures are so different that the results of one cannot be given any interpretation at all from the other point of view.

Speech synthesis

There are several papers on speech synthesis being presented at this Congress, so there is no need for me to make many comments here. We can see that it is now possible to synthesize a wide variety of languages using about 15 parameters, such as those used by Gunnar Fant and his associates in the OVE3 speech synthesizer. Provided that we have sufficiently good acoustic analyses, so that we know the correct values for each parameter, we can synthesize with very high fidelity any human language. For example, consider recordings of the Yoruba phrase [ɛkârɔ̀o ó a dúk͡pɛ́] as spoken by a native speaker of that language, and as synthesized by OVE3. The auditory difference between these two recordings is very small. There is no difficulty in synthesizing the Yoruba [k͡p] sound, despite its complex articulation which involves a velaric ingressive bilabial stop occurring at the same time as a pulmonic egressive velar stop (Ladefoged 1964). Similarly we may compare the Zulu phrase [g͡ʗíɓa isiɔʰúmo] as spoken by a native speaker of Zulu and the same phrase as produced by OVE3. There is a slight difference between the synthesized and the real click sounds, because the synthesizer cannot make a click of sufficient intensity relative to the voiced sounds. In addition the nasal in the last syllable is not correct. As we shall discuss later, nasals are sounds that are easy to describe in articulatory terms, but are very complex from an acoustic point of view. However, the auditory difference between the two recordings is not very great. Synthesis of this quality indicates that speech can be very completely specified in terms of 15 parameters, or less. In other words, we can specify any sound in any language simply by stating the correct values of 15 parameters.

Speech synthesis offers a neat way of distinguishing between the aspects of speech that are specific to the language being described, and the aspects that are characteristic of a particular speaker, or a particular style of speech. The speaker and style dependent characteristics can always be stated as variations of parameters that affect utterances as a whole. Usually they are variations in the mean and range of parameters such as F0, (the pitch) or F1, F2, and F3 (the formant frequencies). All the aspects of speech that are language specific can be stated in terms of rules and a table indicating the relative values of the phonological units used in the rules. Given these rules and the table, one can construct a good simulation of any utterance in the language simply by stating the underlying phonological units that are present. One can also simulate differences between particular speakers by appropriate manipulations of the overall values of the parameters.

These points may be demonstrated by reference to a system for synthesizing tone languages by rule. The input for a Yoruba phrase is shown in Figure 1. Since our computer can handle only capital letters we have to represent the underlying phonemes in terms of the so-called ARPABET symbols shown in line (2). The vowels are represented by two symbols the second of which is always H or X. The number after the vowel is the tone of the syllable. In these rules Yoruba is treated as a language with three phonemic tones. A more conventional IPA transcription is shown in line (3).

The computer accepts the ARPABET symbols, applies the appropriate rules so as to generate the equivalent of a narrow phonetic transcription, looks up the target values for each phone in a stored table, and then calculates the values of the parameters for use by the OVE3 synthesizer. It also uses a set of tone rules which we will discuss later to produce the pitch contour of the phrase, which is as shown in line (1) of Figure 1.

The synthesized phrase is not as lifelike as the Yoruba utterance previously discussed because the rules and the target values are not exactly right. But it is clear that we can specify most of the phonetic characteristics of this Yoruba phrase in this way. Moreover we can make it sound like the same Yoruba phrase pronounced by another speaker simply by altering the ranges over which the parameters are allowed to vary. In the one recording the formant frequencies and the fundamental frequency are approximately 10% higher than in the other. The difference between the two recordings is equivalent to the difference between two speakers with identical accents but different voice qualities. Both are equally good approximations to Yoruba.

Systems for speech synthesis by rule can also be used for checking aspects of the phonological descriptions of languages. For example, some of the tone rules of Yoruba have been formulated by La Velle (personal communication). These rules have been incorporated into the set of rules for synthesizing Yoruba. Figure 2 shows a section of the FORTRAN code that was used in generating the Yoruba utterances that have just been discussed. One does not have to be a computer programmer to see that these FORTRAN statements are directly equivalent to phonological rules. In this section of the code:

SYL	refers to the tone of the syllable in the rule.
THIS	refers to the position of this syllable.
FIRST	implies that THIS syllable is the first in the phrase.
LAST	implies that THIS syllable is the last in the phrase.
PREV	refers to the previously specified tone.
NEXT	refers to the next specified tone.

(1)

(2) OH2RAH3 EX2LEX3DEX1

(3) o r á ɛ l ɛ́ d ɛ̀

Figure 1. Line (1), the intonation contour of a Yoruba phrase, as calculated by a computer, given the input shown in line (2). An IPA transcription of this phrase is given in line (3).

```
  IF (SYL.EQ.HI.AND.THIS.EQ. FIRST)        P =BASE+2*N
  IF (SYL.EQ.MID.AND.THIS.EQ. FIRST)       P =BASE+N
  IF (SYL.EQ.LO.AND.THIS.EQ. FIRST)        P =BASE
  IF (THIS.EQ.FIRST) GOTO 500
  IF (SYL.EQ.MID.AND.PREV.EQ.HI)           P =P-N
  IF (SYL.EQ.HI.AND.PREV.EQ.MID.
/  OR.SYL.EQ.MID.AND.PREV.EQ.LO)           P =P+N
  IF (SYL.EQ.HI.AND.PREV.EQ.LO)            P =P+2*N-D
  IF (SYL.EQ.LO.AND.PREV.EQ.HI)            P =P-2*N
```

Figure 2. Part of the FORTRAN code for calculating the fundamental frequency of the Yoruba phrase shown in Figure 1.

The tone rules can set P, the target pitch of the syllable, in terms of the value of P for the previous syllable, and some combination of the constants N (the interval between tones) and D (the downdrift factor). The constants P,N, and D can be set by the user. The user can also set the BASE or pitch at which a low tone would be if there were a low tone at the beginning of the utterance.

The first three statements in the example set of tone rules in Figure 2 set the pitch for the first syllable in an utterance. These rules are disjunctively ordered with respect to the rest of the rules; they are therefore followed by a GOTO statement that ensures that the next set of statements is bypassed. The remaining statements account for downdrift during the utterance. In addition to these and other phonological rules, the program includes a set of statements for turning the set of target pitches calculated for each syllable into a continuous curve that is stretched over the utterance as a whole.

At least potentially, speech synthesized according to a set of rules can be used to specify all the phonetic (and some of the phonological) characteristics of languages. Linguistic descriptions of this kind fill most of the requirements of scientific theories: they are empirically testable, and they can be expressed in terms of numbers referring to a universal set of parameters. Furthermore, these parameters are well-defined, in general physical terms, and probably constitute the minimum set of parameters that can be used for specifying the phonetic characteristics of languages. The problem is that these descriptions are very hard to relate in their entirety to phonological descriptions of languages that have been proposed by linguists. The reason for this is that a linguist making a phonological description is trying to do much more than specify the phonetic characteristics of languages. Linguists are trying to explain both how and why the sound patterns of languages differ from one another.

Linguistic phonetics

Languages (at least those which we as phoneticians are concerned with) have to be pronounceable. They will, on the whole, tend to have sounds that speakers find relatively easy to say. Complex vocal gestures are likely to be simplified as the language changes.

But languages also have to be readily understandable. Their sounds must be sufficiently distinct in their acoustic correlates for a listener to be able to identify words without difficulty. If the degree of perceptual similarity between two sounds becomes too great, then the language will not be able to communicate facts that rely on the difference between these sounds.

Table I. The features of vowel quality.

Feature	Auditory or physiological	Observable correlates
Height	auditory	frequency of formant one
Back	auditory	difference between frequencies of formant two and formant one
Rounding	physiological	lip position
Wide	physiological	size of the pharynx
Rhotacized	auditory	frequency of formant three
Nasalization	physiological	position of the soft palate

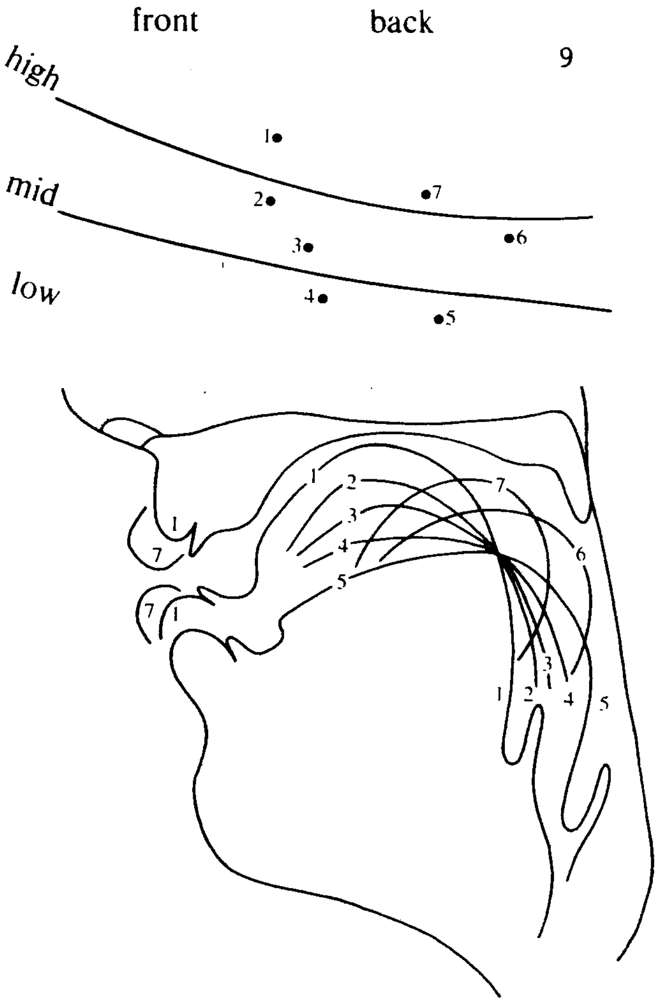


Figure 3. The positions of the tongue during the vowels in my pronunciation of the words: (1) heed, (2) hid, (3) head, (4) had, (5) father, (6) good, (7) food. The highest points of the tongue are shown on a slightly expanded scale in the upper part of the figure. (After Ladefoged, 1975)

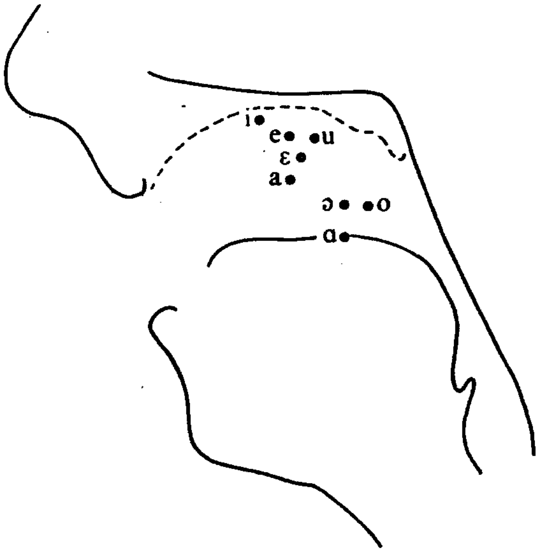


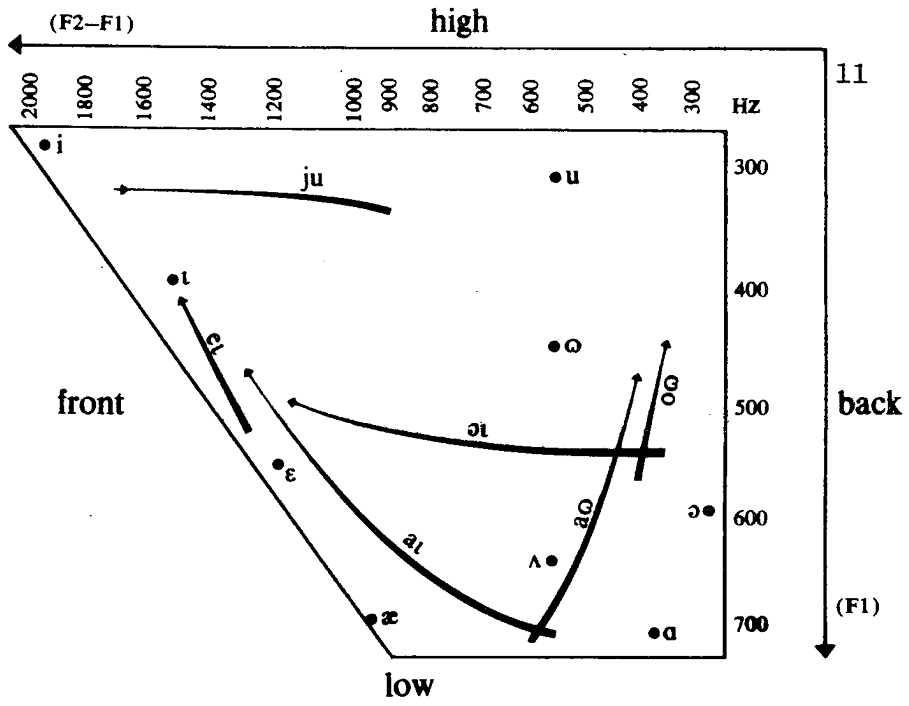
Figure 4. The highest points of the tongue in a set of cardinal vowels pronounced by S. Jones (1929). (After Ladefoged, 1975)

The sound patterns that arise in languages are largely the result of the interplay of these two forces (Martinet, 1955). On the one side there is a tendency towards economy of articulation, which often leads to assimilations. This tendency accounts, for example, for the change in place of articulation whereby "input" becomes [ˈιmpωt]. On the other side there is a tendency towards maximizing auditory contrasts. This tendency leads to the distribution of vowels in the acoustic vowel space as described by Lindblom and his colleagues (Lindblom, forthcoming; Liljenkrants and Lindblom, 1972). Of course there are also other reasons, such as culture contact and prestige factors which cause languages to change. But these two general principles -- ease of articulation and the maximization of perceptual differences -- are the major reasons for the universal tendencies that occur in the sound patterns of the languages of the world. On some occasions in the history of a language the one principle has been more dominant and on other occasions the other. Languages differ because the resolution of the conflict between the two opposing forces has been made in different languages in different ways.

Phonologists who want to explain the sound patterns of languages must refer to both of these principles. This means that they must consider speech sounds sometimes from an articulatory point of view and sometimes from an acoustic point of view. In practice, this involves dividing the sounds of languages into sets of overlapping classes by means of features, some of which are measurable in physiological terms and some in acoustic terms. A possible set of features that can be used in this way has been described elsewhere (Ladefoged 1975). In the present paper I will consider only the subset of features shown in Table I. These are the features that are appropriate for explaining how and why vowels function as they do in various languages.

Traditional descriptions of vowels are still in terms of the position of the highest point of the tongue. They are thus articulatory descriptions. Generative phonologists (Chomsky and Halle, 1968; Hyman, 1975; Anderson, 1975) also refer to the height and degree of backness of the tongue when describing vowels. But for nearly half a century it has been known that in following this practice "phoneticians are thinking in terms of acoustic fact and using physiological phantasy to express the idea" (Russell, 1928). Data based on x-rays showing the tongue positions for some of my English vowels are reproduced in Figure 3. The highest points of the tongue are shown in the upper part of the figure. Obviously some of them are not in very good agreement with traditional descriptions of these vowels. We can only pretend, for example, that the vowel in "food" is higher than the vowel in "hid" by considerably warping the line separating high vowels from non-high vowels. Similarly, as shown in Figure 4, the highest points of the tongue in a set of cardinal vowels (Jones, 1929) are not at all similar to the descriptions

Figure 5. The formant frequencies of American English vowels. (After Ladefoged,1975)



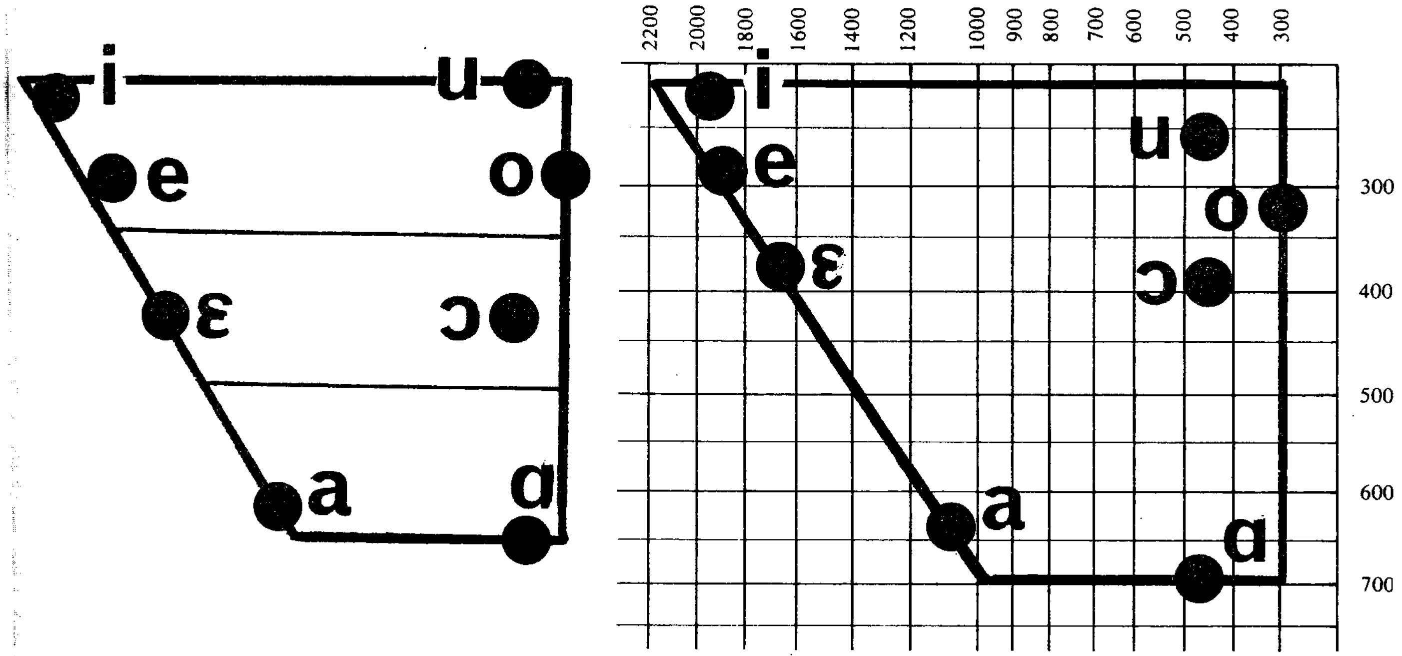


Figure 6. An auditory representation of some Danish vowels (after Uldall, 1933) and a plot of the formant frequencies of these vowels (data from Fischer-Jorgensen, 1972). The frequency of the first formant is shown on the vertical axis and the difference between the first and second formant frequencies is shown on the horizontal axis.

that have been given of these vowels. There is not even a good separation between front vowels and back vowels in this diagram; and the relative heights are wildly wrong.

We can easily see that phoneticians have actually been very accurate in describing vowels by comparing acoustic analyses with traditional descriptions. Figure 5 shows a formant plot of the vowels of American English. The values for the monophthongs are those reported by Peterson and Barney (1956). Those for the diphthongs are as reported by Holbrook and Fairbanks (1962), normalized so as to fit in the same range as the other data. The vertical axis shows the frequency of the first formant, and the horizontal axis the difference between the frequencies of the first two formants. The scale of the vertical axis is twice that of the horizontal axis, with, in both cases, the intervals being proportional to the perceptual distances in mels (Stevens and Volkman, 1940).

When the physical measurements are represented in this way they correspond closely to the auditory judgments of skilled phoneticians. Apart from the scales around the outside, Figure 5 is very similar to diagrams that have been produced by Kenyon and Knott (1953), Jones (1956), Ward (1945), Gimson (1970), and many other phoneticians. There is no doubt that all these authors, each of whom claimed to be describing tongue positions, were actually making fairly accurate judgments of the frequency of the first formant and the difference between the frequencies of the first and second formant.

Examples of this kind could be given from published data on many different languages. Perhaps one of the most striking concerns the vowels of Danish. Figure 6 shows, on the left, some of the vowels of Danish just as they were represented by Uldall (1933) in his *Danish Phonetic Reader*. Plots of the mean formant frequencies of these vowels as determined by Fischer-Jørgensen (1972) are shown on the right of the figure. The similarity between these two sets of data, published nearly 40 years apart, is overwhelming. Notice, for example, how exactly Uldall has captured the correct relationship in the so-called back-front dimension, with [u] and [ɔ] being slightly further forward than [ɑ] and [o]. Compare also the way in which the so-called height dimension is represented, with the distances between [i, e, ɛ, a] steadily increasing, and [ɔ] being almost exactly half way between [u] and [ɑ]. All these correspondences are remarkable testimony to the skill of Uldall as a phonetician.

Based on this and a large amount of similar evidence there is no doubt that the primary phonetic features of vowel quality are the frequency of the first formant, and the difference between the frequencies of the first and second formants. It seems advisable to continue referring to these features as Height and Backness, with the understanding that they should be measured in acoustic, and not in physiological terms. Furthermore, there is no doubt that these features are useful in phonological descriptions of the sound patterns that occur in languages. They have been used in nearly every phonological description that has appeared in the last 100 years. Vowel mutations of many different kinds can be neatly expressed in terms of the space they define. They permit us to explain many points about why languages behave in the ways that they do in terms of perceptual similarity and perceptual contrast.

We might note in passing that each of these features, Height and Backness, is from a phonological point of view a single multivalued feature. As I have discussed elsewhere (Ladefoged 1972a, 1972b) there is nothing to be gained and much to be lost by splitting Height, for example, into [∓ High] and [∓ Low]. Languages work in terms of a single scale of Height.

Because of the arguments advanced above, I have defined the features Height and Back in acoustic terms. This is not to deny the reality of their articulatory correlates. It would be possible to describe Height and Back in terms of tongue positions. But, if we apply these terms literally to tongue shapes, we end up with unnatural classes of vowels that phonologists could not use in describing languages. We find for example, that we have to consider [ɛ] as being as close to [u] as to [ɔ] in terms of tongue height; and we have to put [u] into the class of front vowels. Such classifications are completely inappropriate if we want to describe how languages work. There is a very complicated relation between actual articulatory positions and formant frequencies (Fant, 1960). Descriptions of the features Height and Back could be given in terms of measurable articulatory parameters, but they would have virtually no explanatory power. They would not reveal why vowels pattern as they do.

For some features of vowel quality, however, it is the acoustic correlates that are complex and non-explanatory, while the articulatory correlates are easily measurable and obviously account for the observable sound patterns. This is demonstrably true in the case of the feature Round. As may be observed from the data in Fant (1960), adding lip rounding has different effects for different vowels. In the case of high front vowels, lip rounding causes a lowering of F3, and has virtually no effect on F2. For lower front vowels the reverse is true: F3 is unaffected, but F2 is lowered. For back vowels both F3 and F2 are lowered, in different degrees, depending on the vowel in question. All these diverse acoustic effects have a common, easily measurable, articulatory correlate, the degree of lip rounding. It is this aspect of the vowel that needs to be mentioned in accounting for such phenomena as the rounding of vowels after labial consonants in Turkish. It would be possible to describe the sound pattern of Turkish in terms of rules that accounted for the observable variations in formant frequencies. But rules of this kind would give us no insight into the causes of the phenomena.

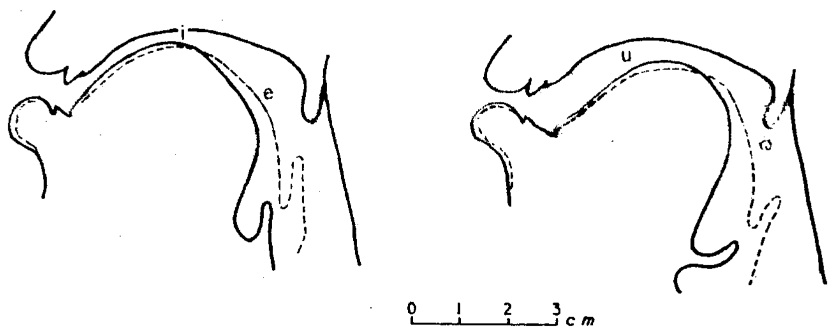


Figure 7. The tongue positions in some Igbo vowels. (Ladefoged,1964).

Table II

A phonetic specification of some of the sounds of English in terms of percentage values of the prime features. The consonants are assumed to be the allophones that occur before the vowel [aɩ] in a stressed syllable. The vowels are assumed to be the allophones that occur at the beginning of a stressed syllable. There is a great deal of guesswork in this table.

	p	t	k	f	θ	s	ʃ	m	n	w	j	r	l	h	i	ɩ	æ	ɔ	oω
Glottalic	60	60	60	50	50	50	50	50	50	50	50	50	50	50	50	50	50	50	50
Velaric	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Voice	0	0	0	0	0	0	0	80	80	80	80	80	80	20	80	80	80	80	80
Aspiration	90	95	100	70	70	70	70	0	0	0	0	0	0	60	0	0	0	0	0
Place	100	85	60	95	90	85	80	100	85	60	70	80	85	60	75	70	60	40	55
Labial	100	5	5	90	5	40	60	100	5	80	5	40	5	5	25	20	5	30	40–80
Stop	100	100	100	90	90	90	90	100	100	80	80	70	70	0	75	70	5	70	70
Nasal	0	0	0	0	0	0	0	100	100	0	0	0	0	5	0	0	5	0	0
Lateral	0	0	0	0	0	0	0	0	0	0	0	0	100	0	0	0	0	0	0
Trill	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Tap	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Sonorant	0	0	0	5	5	15	15	75	75	70	70	75	80	5	80	85	95	95	95
Sibilant	0	20	0	10	20	100	90	0	0	0	0	0	0	10	5	0	0	0	0
Grave	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—
Height	100	100	100	100	100	100	100	95	95	90	90	90	90	5	85	60	5	20	30–50
Back	—	—	—	—	—	—	—	—	—	90	0	—	—	40	0	20	50	100	90–95
Round	0	0	0	0	0	40	60	0	0	90	0	60	0	0	0	0	0	40	40–90
Wide	50	50	50	50	50	50	95	50	50	50	95	30	50	50	95	60	50	20	60
Rhotacized	0	0	0	0	0	0	0	0	0	15	0	100	0	10	0	5	10	15	15
Syllabic	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—

(from Ladefoged, 1975)

As far as the principle properties of vowels are concerned, it seems as if languages work in terms of two acoustic properties, Height and Back, and one articulatory property, Round. As we have seen, the sound patterns that arise can be most easily explained in these terms, and phoneticians judging vowel quality seem to behave in a corresponding way. It is interesting to note that phoneticians can make fairly accurate judgments of the features Height and Back, even when working from recordings. But most phoneticians are not so proficient at assessing the degree of lip rounding when they cannot see the speaker (Ladefoged, 1967; 132-42).

The phonologies of many of the languages of the world can be described by reference to no more than the three features Height, Back and Round. But some languages cannot be adequately described without reference to additional properties. There are several instances of vowel harmony languages (see review of the literature in Lindau, 1975) where the vowels have to be divided into two sets. In the older literature it was often said that the vowels in the one set were tense, and in the other set lax. These labels, however, were seldom defined in terms of quantifiable properties.

In the last decade, measurable physiological data on vowel harmony languages has become available. Figure 7 shows x-ray tracings of the vowels of Igbo, which is one of the languages where there are two phonologically distinct sets of vowels of this kind. These x-rays, which were published some years ago (Ladefoged, 1964) led to discussion of the possibility of distinguishing between the two groups of vowels in terms of the degree of advancement of the root of the tongue (Stewart, 1967; Pike, 1968). This possibility has now been amply confirmed in a number of languages, with consequences that are discussed by Lindau (forthcoming). There is no doubt that the vowels of many languages have to be specified in terms of a feature Wide (or Expanded). Like Round, the acoustic consequences of this feature are complex, but the physiological correlate can be specified in terms of a single scale, in this case the size of the pharynx.

There is a small number of languages in which vowels have to be specified in terms of the feature Rhotacized. This is the feature that specifies what is usually called r-colouring. Many speakers of American English have some r-coloured vowel in words such as "sir." This is one of the most clear cut examples of a sound for which there are two different possible articulatory gestures that have virtually identical auditory consequences (Uldall, 1958). This sound can be produced with the tip of the tongue raised; or it can be made with the tip lowered and with the tongue bunched up lengthwise and slightly retracted. Because there are two different ways of making this sound there is no point in trying to specify or explain the effects of rhotacization in articulatory terms. It is much more appropriate to quantify the feature Rhotacized in terms of the degree of lowering of the third formant, which is the invariable acoustic correlate.

This kind of specification also helps us to understand why rhotacized front vowels often appear to be centralised. The tongue may or may not be retracted, depending on the speaker. But the lowering of the third formant and the concomitant lowering of the second formant make these vowels have a higher value of the feature Back, irrespective of the articulatory action involved.

The final phonological feature that is needed in the specification of vowels is Nasal. Many languages (eg. French, Ewe and Chinatec) distinguish between oral and nasal vowels. From an acoustic point of view, the correlates of nasality are extremely complicated; there is no single acoustic property that can be measured. In fact, I do not know if anyone has been able to devise even a complex algorithm that will enable one to quantify acoustically the degree of nasality in a vowel. I doubt whether one can tell invariably, on the basis of the acoustic data alone, whether a vowel is nasalized or not. But the articulatory correlate, the degree of velic opening, is fairly straightforward. Nasalized vowels clearly belong together as a class because of their physiological properties.

In summary, of the six phonological features we have been discussing, three (Height, Back and Rhotacized) are acoustically based, and three (Round, Wide and Nasal) have an articulatory basis. Each of the acoustic features is associated with particular articulatory gestures, and each of the articulatory features has certain acoustic effects. But the definition of each feature has to be in either acoustic or articulatory terms (and not both) if it is to have any explanatory power. An explanation of a event is an account of that event in terms of something else. A good explanation usually involves describing an event in terms of concepts that are empirically verifiable in another discipline. If our explanations of phonological events are to be satisfactory they must refer to non-linguistic concepts that are known to be governed by more general laws.

Phonetic specifications

We are now in a position to make an overall comparison of the two systems of specifying languages: - acoustic speech synthesis by rule and linguistic specifications in terms of generative phonology. Note that, at least potentially, descriptions made in either of these ways form equally viable scientific theories. They both can be reduced to empirically testable propositions expressed in valid, reliable, and significant numbers; and they both explain something (different in each case) about naturally observable phenomena.

Speech synthesis by rule is obviously a completely quantified system. Its explanatory role concerns the relation between the physical limits of different communications systems, and the rate at which linguistic and paralinguistic information can be conveyed. It helps us answer questions such as why are individual voices recognizable to a limited extent, and why are there no machines (or computer systems) that will recognize sentences that may use a vocabulary of more than 1,000 words.

Phonological features may also be quantified. Each of them is defined in terms of a physical scale, so that at the systematic phonetic level each of them specifies a measureable physical variable. Table II shows a complete set of specifications for a number of English allophones. The box in the lower right hand corner of the table marks off the set of values for specifying vowels using the features that have been discussed in this paper. Definitions of the other features have been given elsewhere (Ladefoged, 1975).

It should be noted that many of the values given in Table II are only educated guesses. All the values have been rounded off to the nearest 5%, because it seems unwise to pretend to a more accurate specification. In the case of the vowels, it is possible to see how some of the values have been calculated by reference to the data reported in this paper. The values for Height and Back reflect the measured formant frequencies shown in Figure 8. Thus [ɩ] has a Height value of 60%, because (for the average Peterson-Barney subject) it has a first formant frequency of 400 Hz. The range of first formant frequency for the subject is taken to be from a maximum of a little over 700 Hz to a minimum of about 200 Hz, a total of slightly more than 500 Hz. The actual value in this vowel is 60% of the distance from the maximum value to the minimum. Similarly the Back value of 20% indicates that it is 20% of the distance from the maximum possible difference between the formants to the minimum possible difference. The values for the feature Round are based on measurements reported by Fromkin (1964); and those for the feature Wide are based on my own measurements of x-ray tracings.

The values are relatively well determined for all the features that are predominantly concerned with vowels. But let me emphasize again that the values for many of the other features are merely the best estimates that can be made at this time. They are listed simply to make the point that each feature is potentially measureable for every sound.

The values given in Table II are the percentages of the full range of each feature. These are all that are significant from a linguistic point of view. In a particular utterance, the actual values, (in Hz or in msec or in whatever unit is required) will depend on nonlinguistic factors such as the size of the speaker's head. All that is linguistically significant is that the language being described has, for example, a vowel [ɩ] in which the first formant is 60% of the full range, as opposed to the similar vowel in another language in which the value may be 55%.

The notions expressed in the preceding paragraphs are completely in accord with the standard theory of generative phonology. Chomsky and Halle (1968) have also said that each feature represents a physical scale, and that at the systematic phonetic level, sounds can be specified in terms of numerical values. But in practice generative phonological descriptions almost never express anything in numbers. A systematic phonetic specification of the kind shown in Table II is very rare. In fact, I have never seen another description of English that gives even estimated numbers for the values of a complete set of features.

It is interesting to consider why generative phonologists have given lip service to the notion that features represent physical scales, but never give systematic phonetic descriptions of sounds in terms of these scales. The answer probably lies in the standard generative view of the nature of language. Thus in Chomsky's (1965) view, a description of a language is an account of the linguistic competence of a native speaker of that language. Unfortunately, this notion does not readily lend itself to quantifiable descriptions.

I do not want to give the impression that I consider the goals of the standard generative theory of linguistics to be wrong, or even inappropriate. But I do feel that it should be realized that attempts to describe the linguistic competence of a speaker form only one way of trying to describe a language. A language has an existence quite separate from its existence within the competence of any of its speakers. A language is a social institution, just like any code of behaviour belonging to a social group. It is, for example, possible to describe the linguistic behaviour of a particular group in much the same way as it is possible to describe their moral behaviour. Clearly codes of moral behaviour exist; and it is possible to think of them as existing within the competence of an individual - whose performance may not always match his competence in this respect. Descriptions of moral behaviour in terms of an individual's moral competence are quite common. In these descriptions an individual's moral competence is often referred to as his conscience. But sociologists, anthropologists, and other social scientists usually avoid terms such as conscience, and use other ways of describing moral behaviour. They are apt, for example, to describe and explain it in terms of forces, such as the conflict between the acts that benefit society and those that benefit the individual. Just as languages differ because they resolve the conflicting forces acting on them in different ways, similarly societies differ because they resolve the conflict between the good of society and the liberty of the individual in different ways, and hence have different codes of moral behaviour.

If we as phoneticians want to explain things about languages, I would suggest that we should behave more like sociologists, and less like moralists. I do not deny the value of the notion competence; it seems to me to be exactly equivalent to one's linguistic conscience. But languages exist, just like a code of moral behaviour exists, outside any individual. For many purposes descriptions in terms of a speaker's competence are simply inadequate. Explanations of linguistic phenomena must reflect the fact that a language cannot be the property of a single individual. A language must be something that is shared by the members of a social group. For any real linguistic communication there must be at least two individuals, a speaker and a listener, each of whom has different needs with regard to the language. Remembering this will affect the way we choose to describe phonetic events.

Phonetics ,as Abercrombie (1967) has pointed out, is the study of the medium of spoken language. Part of our task as phoneticians is to describe each language in terms of the properties of the medium. We can do this simply by describing one aspect of the medium -- say the acoustic aspect -- in the least redundant way possible, using a minimum set of parameters. Or we can describe the language in a way that reveals all the phonological patterns that occur. As I have tried to show, these different tasks are not easy to reconcile. But whichever we undertake, let us hope that we tackle it in ways that Lord Kelvin would appreciate.

References

Abercrombie, D.A. (1967) *Elements of General Phonetics,* Edinburgh: Edinburgh University Press.

Anderson, S. (1975) *The Organization of Phonology,* New York: Academic Press.

Chomsky, N. (1965) *Aspects of the Theory of Syntax,* Cambridge, Mass: M.I.T. Press.

Chomsky, N. and Halle, M. (1968) *The Sound Pattern of English,* New York: Harper and Row.

Fant, C.G.M. (1960) *Acoustic Theory of Speech Production,* The Hague: Mouton.

Fischer-Jørgensen, E. (1972) "Formant frequencies of long and short Danish vowels" in *Studies for Einar Haugen* (ed. E.S. Firchow et al.) The Hague: Mouton.

Fromkin, V. (1964) "Lip positions in American English vowels." *Language and Speech* 7, 215-225.

Gimson, A.C. (1970) *An Introduction to the Pronunciation of English,* London: Edward Arnold.

Holbrook, A. and Fairbanks, G. (1962) "Diphthong formants and their movements" *J. Speech and Hearing Research 5,* 38-58.

Hyman, L. (1975) *Phonology: Theory and Analysis,* New York: Holt, Rinehart and Winston.

Jones, D. (1956) *An Outline of English Phonetics,* Cambridge: Heffer.

Jones, S. (1929) "Radiography and pronunciation" *British J. Radiology* 3, 149-150.

Kenyon, J.S. and Knott, T.A. (1953) *A Pronouncing Dictionary of American English,* Springfield, Mass: Merriam.

Ladefoged, P. (1964) *A Phonetic Study of West African Languages,* Cambridge: Cambridge University Press.

Ladefoged, P. (1967) *Three Areas of Experimental Phonetics,* London: Oxford University Press.

Ladefoged, P. (1972a) "Phonetic prerequisites for a distinctive feature theory" in *Papers in linguistics and phonetics in memory of Pierre Delattre* (ed. A. Valdman). The Hague: Mouton.

Ladefoged, P. (1972b) "Phonological features and their phonetic correlates" *J. International Phonetic Association 2,* 2-12.

Ladefoged, P. (1975) *A Course in Phonetics,* Harcourt Brace Jovanovich.

Liljenkrants, J. and Lindblom, B. (1972) "Numerical simulation of vowel quality systems: the role of perceptual contrast" *Language* 48, 839-62.

Lindau, M. (1975) "Vowel features" *Working Papers,* Phonetics Laboratory, Lund University.

Lindau, M. (forthcoming) " Features for Vowels " *UCLA Working Papers in Phonetics* 30.

Lindblom, B. (forthcoming) "Experiments in sound structure" (This Congress).

Martinet, A. (1955) *Economie des changements phonétiques,* Berne: A. Francke.

Peterson, G.E. and Barney, H.L. (1956) "Control methods used in a study of the vowels" *J. Acoustical Society of America* 24, 175-184.

Pike, K. (1968) "Tongue root position in practical phonetics" *Phonetica* 17, 129-40.

Russell, G.O. (1928) *The Vowel,* Columbus: Ohio State University Press.

Stevens, S. and Volkman, J. (1940) "The relation of pitch to frequency - a revised scale" *American J. Psychology* 53, 329-38.

Stewart, J.M. (1967) "Tongue root position in Akan vowel harmony" *Phonetica* 16, 185-204.

Thompson, W. [afterwards Lord Kelvin] (1889) *Popular lectures and addresses,* London: Macmillan.

Uldall, H.J. (1933) *A Danish Phonetic Reader,* London: University of London Press.

Uldall, E.T. (1958) "American 'molar' r and 'flapped' t" *Revista do Laboratorio de Fonetica Experimental, Coimbra 4,* 103-106.

Ward, I. (1945) *The Phonetics of English,* Cambridge: Heffer.

The Wheres and Whyfores of Preach Seduction

Victoria A. Fromkin

[Paper presented at the 90th Meeting of the Acoustical
Society of America, San Francisco, California,
November, 1975]

Denis and Pinson in their now classic work *The Speech Chain* discuss "what happens (when) two people talk...to each other" and state that "The first thing the speaker has to do is arrange his thoughts, decide what he wants to say and put what he wants to say into linguistic form." They go on: "The message is put into linguistic form by selecting the right words and phrases to express its meaning, and by placing these words in the correct order required by the grammatical rules of the language." It is twelve years since the publication of this statement and we still do not know the sequence of events or stages, *how* the message is put into linguistic form, *how* the right words and phrases are selected and ordered, or even *how* the words and rules are stored in the brain. But we are beginning to arrive at some hypotheses. An increasing number of researchers are turning to speech errors, i.e. deviations from the intended utterance or message, to look for the answers to these important questions.

The more we look at deviant speech the more we recognize the incredible comlexity of the mechanisms involved. We find for example that the speaker does not always select "the right words...to express (the) meaning" of his message. But, the words selected, even when substitutions are made, seem to be highly constrained. This paper will discuss a limited class of errors in the attempt to reveal what factors must be considered in constructing a model of the speech chain. Unfortunately it will present more questions than answers.

Speech errors do support the suggestion that the speaker must first "arrange his thoughts" before the message is put into linguistic form. There is no apriori reason why this should be so. One could instead propose that the message is originally formed in linguistic units, e.g. words. Hill (1972) suggests that "a preexistent message which is encoded into language is an entity not necessary for a coherent picture of sentence formation."

Consider however the word substitution errors given in your handout under (1). The intended utterance appears on the left of the arrow; the actual utterance on the right. Dashes, -- , represent a pause.)

(1) a. The Mafia moved into Boston --> The Mafia moved into Italy -- I mean Boston
b. I'm going to England in May --> I'm going to April in May
c. It will be about 3:10 when we get to the airport. --> ...when we land -- get to...
d. I thought Westerns were where people ride horses instead of cars --> ...ride horses instead of cows
e. I better give you a map --> ...give you a calendar
f. He's got a wedding band on his finger --> ...on his wrist
g. are my tires touching the curb? --> are my legs touching the curb?
h. I finally went to the dentist when my gums started to bleed --> ...when my tongues started to bleed
i. that's a horse of another color --> ...of another race
j. put your hands over your ears --> put your eyes over your ears
k. 'Jack' is the subject of the sentence --> ...is the president of the sentence
l. look at the lady with the dachshund --> ... with the Volkswagon

If we form our sentences or utterances initially with linguistic items, these erros can only be accounted for by saying we were unsure of the idea we wished to get across, that is, that there was an ideation error. Such 'confusion' would fail to explain the kinds of errors which are given here. Note that in the cases under (1) the target word and the substituted word have some semantic similarity. In (1)a. the error can be accounted for by positing that some of the semantic properties of Mafia persevere to produce Italy. (Unfortunately the 'state of the art' does not permit a more rigorous definition of semantic similarity). In (1)b. a semantic feature is anticipated or influences and distorts the intended message. Perseveration and anticipation of distinctive features, phonemic segments, morphemes and words also occur in other types of speech errors; thus these word substitutions seem to be produced by similar mechanisms. I have no good explanation for the actual mechanisms involved. Earlier I suggested that the stored lexicon must be arranged according to semantic classes, and that in the translation process from message to linguistic form, that is, in the process of selecting the words to express our thoughts, we may select the wrong item from the right semantic class. This is of course too simple and does not explain the perseveration or anticipation of semantic features. Furthermore, as will be discussed shortly, there are often many factors which seem to give rise to such errors.

The errors under (2) are somewhat different and require further explanation.

(2) a. Nick watered and Sue weeded the garden. --> ...and sea weeded the garden
b. deep phrase marker --> deep freeze marker
c. It spread like wild fire --> ...like wild flower
d. sesame seed crackers --> Sesame Street crackers
e. I see you're not wearing your swim suit --> ...your swimming pool
f. lemons are used by every housewife --> ...by every house word
g. I do better when I eat at the Whole Earth Cafeteria --> ...at the Whole Earth Catalogue
h. chamber music --> chamber maid

Some of these errors can be accounted for as due simply to phonological factors. (2)a. may be merely due to vowel anticipation, and (2)b. to vowel perseveration. In each of these examples the error results in a noun compound or a complex lexical item which is probably stored as a whole in our internal dictionaries. Note further that in a number of these cases there is a formal or phonological similarity between the target word and the substituted word. It is possible therefore that the fact that Sesame Street, for example, is a stored noun compound and is in addition phonologically similar to sesame seed 'conspired' to produce the error.

Hopefully by the next ASA meeting I will have a quantitative analysis of the UCLA corpus of speech errors. The computer program being written by Georgette Silva is now being completed and the processing by Linda Galloway is about to begin. This will provide information on the percentage of phonological errors which result in actual words.

In addition to the errors in (2) in which the substitutions result in the 'selection' or substitution of stored noun compounds there are a large number of errors such as those in (3).

(3) a. a student who has just completed the MA exam --> ...just competed the MA exam
b. I urgently requrest that you release the hostages unharmed --> ...the hostages unarmed
c. What's predictable by rule is part of competence --> ...is part of confidence
d. The detective who issued the warrant --> The defective...
e. I want the number of the Del Amo Fashion Center --> ...Del Amo passion center
f. I'm going to die young, but I'll die less young --> ...I'll die yes lung

Again we can account for such errors simply in phonological terms. (3)a. and (3)b. show a deletion of a segment; (3)c. a change of the stop [p] to the fricative [f] in anticipation of the fricative [s]; (3)d. the anticipation of the place and manner of articulation features of [v] with no change in voicing of the [t] target; (3)e. the substitution of the feature [stop] for the feature [fricative] and (3)f. a typical spoonerism with the transposition of two consonants.

We might wish to ignore the fact that the young man who asked the information operator for the telephone number of the Del Amo Fashion Center was trying to reach the girl he had just moved in with, or that the man who uttered (3)f. had just given up smoking. Perhaps all we are justified in analyzing are the 'hard facts'. But the errors in (4) suggest that we cannot so easily dismiss the effects of 'intruding' thoughts which are either internally or externally engendered.

(4) a. he made headlines --> he made hairlines
b. if someone steals my proto Miwok you and David are automatic suspects --> ...are automatic suffixes
c. the conquest of Peru --> the conquest of Purdue
d. blue sky and yellow sun --> blue sky and blond hair
e. I've never heard of classes on Good Friday --> ...on April 9th
f. the Art of the Fugue --> the arg of the flute
g. she's marked with a big scarlet A --> ...with a big scarlet R

(4)a. was spoken by someone referring to a barber. (4)b. makes more sense when one knows that proto-Miwok is a language with a complex suffixal system. (4)c. was said by Arthur House while he was at Purdue. Note that in these three substituted words, the target words and the substitutions all begin with the same consonant or first syllable. (4)d. could have been influenced by the phonoligical similarity of 'blue eyes' to 'blue sky' which triggered an associat on with blond hair, and/or the semantic similarity of blond with yellow producing the first substituted word which then produced the compound blond hair. (4)e. was uttered when Good Friday fell on April 9th. (4)f. is more complex. It was spoken by a flute player who was to play the flute in the performance. Art of the Fugue must have been phonologically represented in memory prior to the substitution of the word flute (or both must have been present) because of the change of art to arg. Again the phonological similarity of the words 'fugue' and 'flute' may have had a contaminating influence. (4)g. again shows semantic contamination in that scarlet and red are synonyms resulting in the substitution of R, the initial of red, for the intended A, the initial of adultery.

These errors strongly suggest that a model of speech production must include a complex set of mechanisms with many feedback loops and semantic and phonological associateve pathways. In addition, ideas external to the intended utterance may comtaminate or distort the target message.

Any viable model must formally and explicitly provide for the mechanisms by which both ideational and formal associations effect the verbal output. I have no model to propose which can explicate these mechanisms but hopefully in time we will have such a model.

References:

Denes, Peter B. and Elliot N. Pinson (1973) *The Speech Chain: The Physics and Biology of Spoken Language* Anchor Books

Hill, A. A. (1972) 'A Theory of Speech Errors' in Firchow et al (eds) *Studies Offered to Einar Haugen* Mouton

On the relationship between formant bandwidths and vocal tract shape features

Lloyd Rice and Sven Öhman

[Paper presented at the 86th Meeting of the Acoustical Society of America, November 1973]

Theoretical and experimental studies by a number of researchers have shown that the formant frequencies of vocalic speech sounds do not uniquely determine the area function of the vocal tract. In fact, it is possible to change a given area function in infinitely many ways without affecting any of the formant frequencies. If the starting point for such changes is a tube of constant cross-section, and with a low impedance termination at the lips, then a small change which is lengthwise symmetrical about the center of the tract will not affect formant frequencies.

However, there is an acoustic correlate of these changes. We find the formant bandwidths changing inversely to the odd-numbered sinusoidal components of the area function change. Figure 1 illustrates this relationship.

Symmetrical area functions were generated and acoustic waveforms were produced for each area function using a software vocal tract synthesizer similar to that described previously (Rice, 1971). The driving impulse to the vocal tract model was a single spike of one sample pulse width, giving a flat input spectrum. Fast Fourier Transforms of these waveforms gave smooth spectra with well defined resonance curves up through the 4th formant. These spectra were hand fitted using an interactive analysis-by-synthesis technique to determine the frequencies and bandwidths to within two or three Hz accuracy.

In figure 1 the area is shown changing according to the sine of π times the distance along the tube. Three area functions are shown, and I have darkened in the outlines of the tube for position "a" to clarify the picture. The extremes of change shown in positions "a" and "c" represent a peak deviation of 40% of the original cross-sectional area. The change

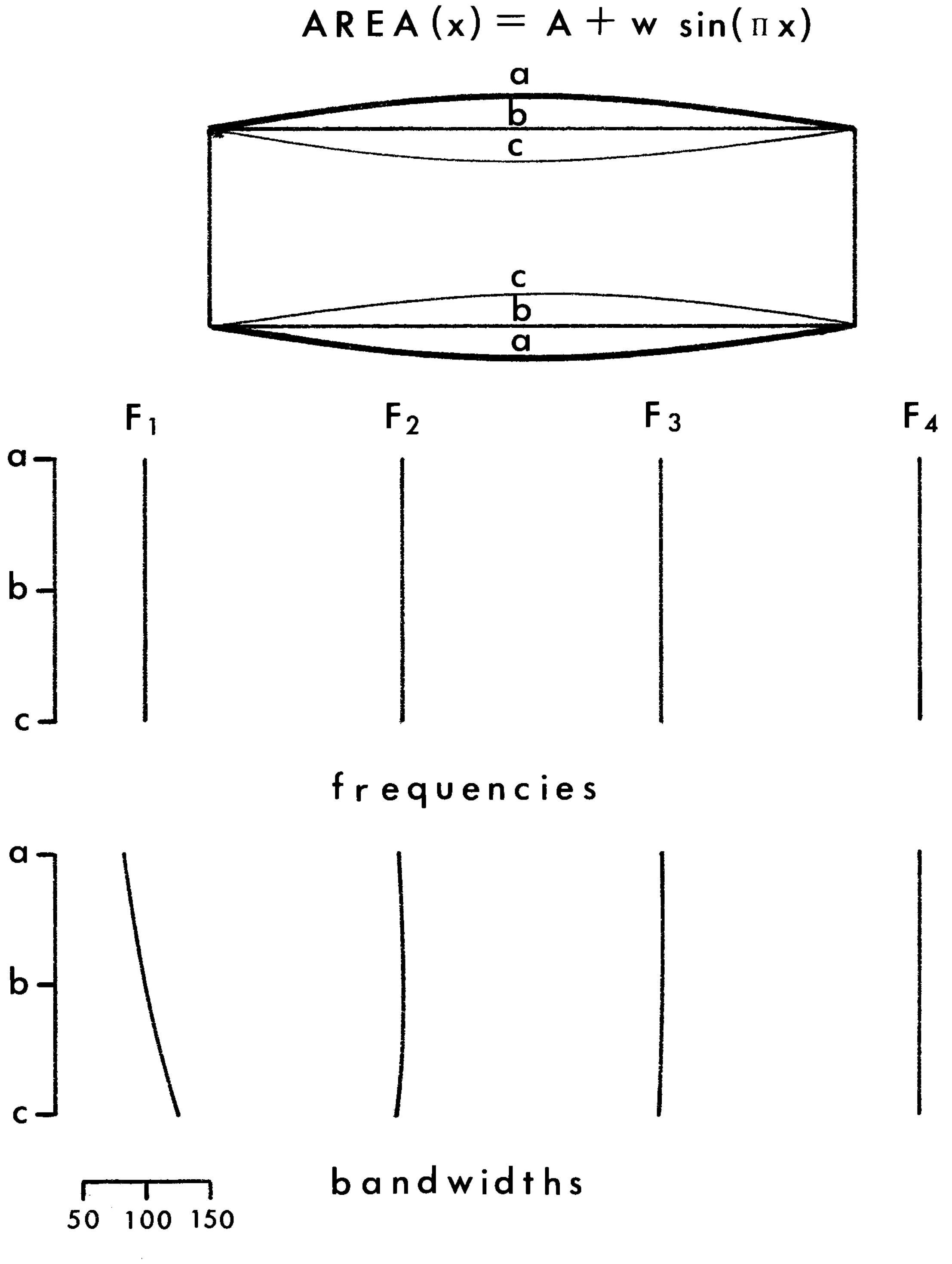

AREA(x) = A + w sin(πx)
a
b
c
c
b
a
F1
F2
F3
F4
a
b
c
frequencies
a
b
c
bandwidths
50
100
150

FIGURE 1

is symmetrical about the longitudinal center line of the tract and, as shown, does not change the formant frequencies. The bandwidth of formant 1 increases from about 80 Hz up to about 125 Hz as the area changes from position "a" to position "c". Notice that bandwidths of other than the first formant are affected only slightly at fairly large deviations from the uniform tube.

In figure 2 is shown an area function perturbation given by the sine of 3 π times the distance along the tube, again a symmetric perturbation, and again we see that the formant frequencies are not changing for small deviations from the uniform tube. The bandwidths here are essentially constant except for the bandwidth of formant 2, which increases sharply as the area changes from "a", through the uniform tube position at "b", to the final position at "c".

Proceeding in a similar manner, we computed the bandwidth effects for the next two sinusoidal components, affecting formants 3 and 4. We have not presented diagrams of these perturbations, because the changes they represent are less natural for the vocal geometry and so are less important to articulatory reality. Instead, we have summarized in figure 3 the results of the bandwidth calculations for the first four formants. Again we see that the bandwidth corresponding to the odd-numbered sine components increases while all others are essentially constant for small perturbations.

After making these figures, we began to wonder how the bandwidths would behave if the perturbations were corrected so as to keep the frequencies exactly at the starting positions. This was accomplished using the perturbation theory model to compute a correction in the area function at each small step in the deviation from the starting configuration. Any time a formant started to move a bit, a perturbation was immediately computed to restore that frequency, but with no constraint being imposed on any bandwidth value. It was determined that the bandwidths behaved in essentially the same patterns as shown here.

$$AREA(x) = A + w \sin(3\pi x)$$

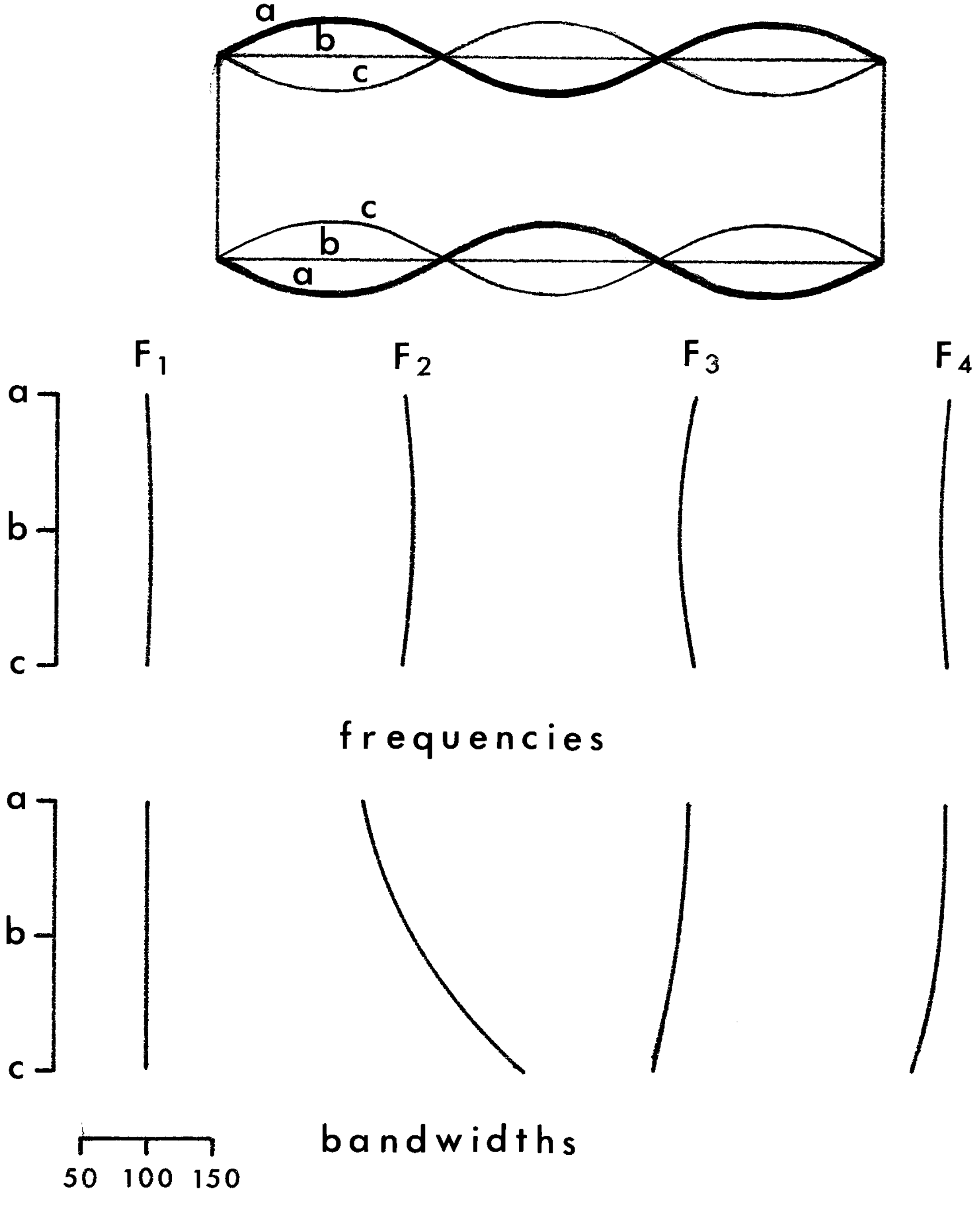


FIGURE 2

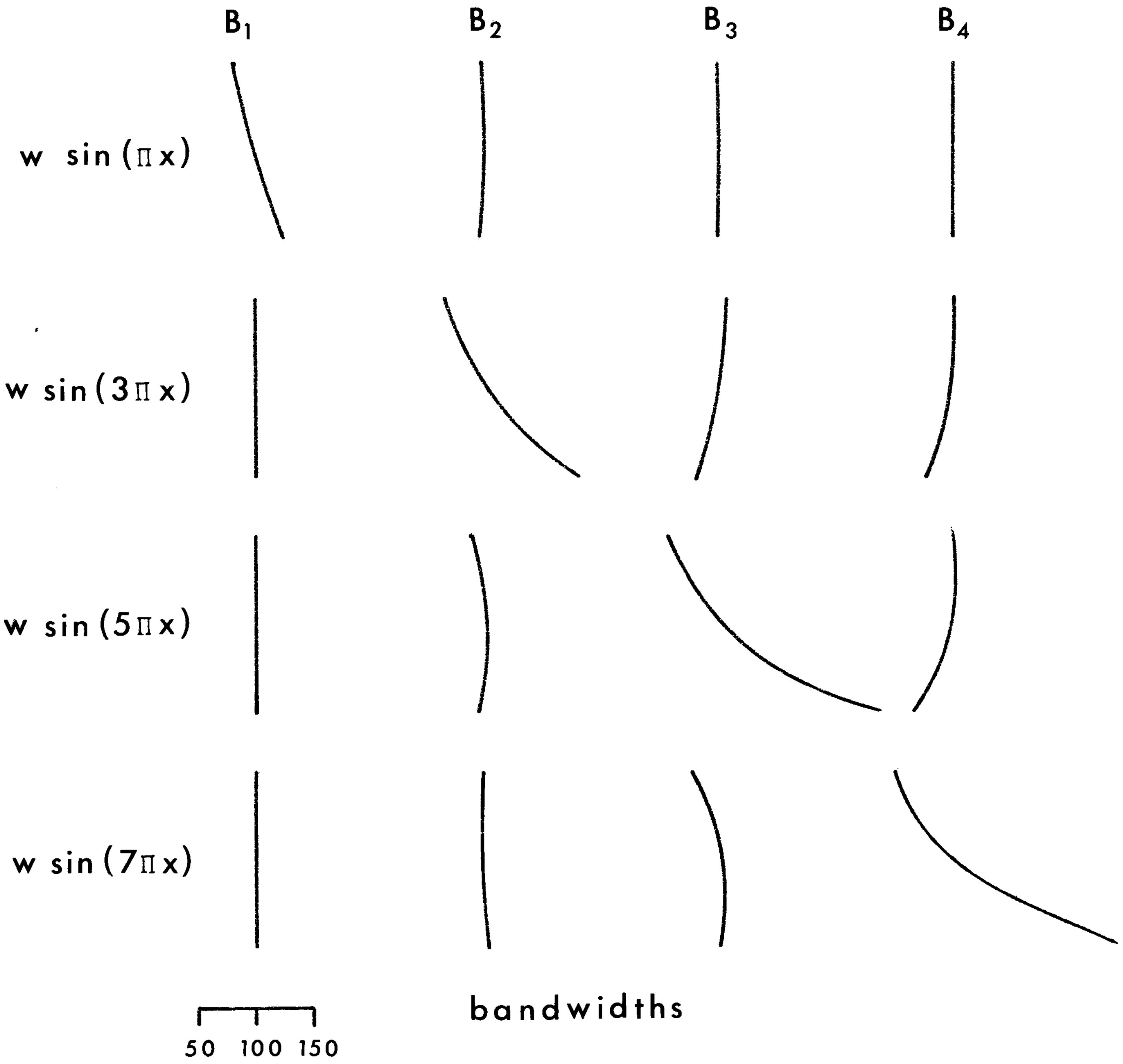

B1
B2
B3
B4
w sin (Πx)
w sin (3Πx)
w sin (5Πx)
w sin (7Πx)
50
100
150
bandwidths

FIGURE 3

Articulatory Tracking of the Acoustic Speech Signal

D. Lloyd Rice

[Published in
Proceedings of Speech Communications Seminar
Stockholm, August 1974, Vol. 1]

Abstract

This paper describes a research project with the goal of tracking a speech signal with a set of articulatory parameters, representing the controls needed to operate an articulatory speech synthesizer. Several major parts of the system are discussed in detail, with emphasis on problems and proposed alternatives. Some preliminary results are presented, using the system to match isolated vowels of a single speaker. Some speaker-dependent characteristics of the system are discussed. The system is expected to be useful for studying the organization of the human vocal tract and may have applications to automatic speech recognition.

Introduction

The Articulatory Tracking system, ARTRAK, is based on ALASS, an automatically controlled speech synthesizer described earlier (Rice 1971). Figure 1 shows a basic block diagram of the ALASS system for matching vowel articulations. At the upper left a test vowel is input and a spectrum analysis is performed. The spectrum is stored for later reference by the comparator. Now the tongue control loop is started. Beginning with an arbitrary initial tongue shape, a vowel is produced by the line analog synthesizer. This vowel is then spectrum analyzed and its spectrum is compared with the stored target spectrum. The comparator determines an error value which depends on the degree of difference between the two spectra. This error value is used by the parameter adjustment strategy in determining which way to readjust the synthesizer so as to minimize the difference between the two spectra. The adjustment process iterates until no more improvement can be made in the spectrum match.

A major problem in designing such a system is that of describing a system of constraints on possible tongue movements. In a recent experiment, Harshman et al (1973) analyzed the tongue shapes in the mid-sagittal plane occurring in the vowels in the words "heed", "hid", "hayed", "head", "had", "hod", "hawed", "hoed", "hood", and "who'd" as spoken by 5 American speakers. Using a PARAFAC (Harshman 1970) analysis system, they showed that the movements of the tongue in a vowel can be described in terms of 2 parameters. The 2 tongue shape parameters determined in this experiment have been incorporated in the ARTRAK model. A third parameter has been added which controls only the lip closure and does not affect the shape of the tongue itself.

Tracking continuous speech

The input signal is a speech waveform which is spectrum analyzed via a Markel autocorrelation LPC filter. In the continuous speech tracking system, successive LPC analyses will be performed at fairly small time steps, depending on the difficulties encountered with jumping out of track. There is a trade-off between minimal spectrum analysis vs. efficient operation of the tracking algorithm. The parameter-controlled synthesis system produces sample spectra for comparison with the stored target spectrum using a difference-minimization algorithm. The synthesizer parameters are automatically adjusted by the minimization algorithm until the best match is found to the target spectrum and the current synthesis parameter settings are recorded as representative of the input signal at the current time frame. A number of details of this matching process will be discussed at some length shortly.

As the analysis moves on in time along the input signal, a good prediction can be made of the approximate parameter values needed based on the previous settings and their motion. Several variations are possible here, including a second-order filtering process already built in the optimization procedure, which could easily be extended to monitor the tracking process. A more interesting question must be considered here, however. That is that the tracking process must have a fair degree of stiffness, or low pass filtering, to prevent the vowel parameters from being thrown out of lock by rapid changes involved in consonant articulations. The research question here is whether filtering threshholds can be found which will allow the tracker to follow rapid formant transitions while enabling it to hold still or move only slowly during voiceless intervals when little or no formant data is available from the input signal.

Spectrum distance measure

One of the most critical components of the optimization procedure is the spectrum comparison method. An error measure must be taken which captures the distance between the synthetic signal and the stored target spectrum in a meaningful way. This requires, in general, that the error measure decrease monotonically as the synthetic spectrum converges on the target from any position in the articulatory space. The approach used in the original system was to compute the sum of squared differences between synthetic and target frequency-differenced spectra as follows:

$$\varepsilon = \sum_{1}^{N-1}((Tn+1 - Tn) - (Sn+1 - Sn))^2$$

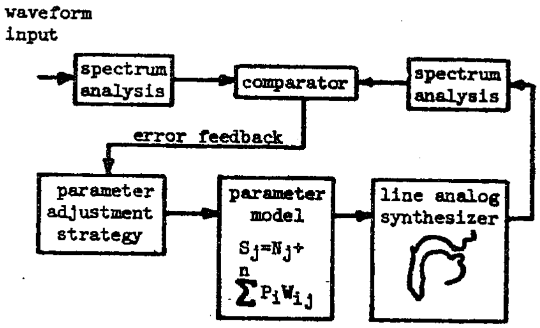
where Tn is the nth point in the target spectrum

Sn is the nth point in the synthetic spectrum.

This was modified slightly in a later version where the analysis of the synthesizer output was performed at twice the sample rate as that of the natural speech because of aliasing problems in desampling the synthesizer output. The line analog synthesizer must run at a specific sample rate depending on the number of sections used (18 sections in the current model force a basic sample rate of 17776 Hz). However, the basic squared difference error measure was the same as described above. Frequency differencing was done to minimize differences due to overall spectrum slope and maximize formant frequency and bandwidth differences. More recent work has shown that this error measure is inadequate over large distances. That is, if the matching procedure is started in approximately the correct region so that the spectrum is not too different from the desired target, the system is able to make changes in the appropriate direction and correctly match the target. If, on the other hand, the procedure is started with parameter values such that the formant peaks are shifted by more than about 10 times the bandwidth of the peak, this error measure does not decrease properly as the synthetic and target formant frequencies move relative to one another. Other error measures are being considered which will incorporate some direct measure of the distances between formants. This involves, of course, the usual formant tracking problem of determining which peak is which, although there now exist in the literature a number of solutions to this problem which work fairly well based on the smooth spectra of all-pole LPC analysis. Alternatively, a more elaborate search procedure is being considered based on a full spectral range error value for movement on each parameter axis, rather than having to optimize from a single scaler error measure.

Shortcut using the Wakita conversion

The present model uses a lengthy computation to determine the the frequency spectrum corresponding to a particular set of values of the tongue control parameters. First the parameter weights and values are multiplied and summed across all parameters to obtain deflections from the neutral shape. These deflections are added to the neutral to obtain the sagittal plane cross-dimension values which are converted to area values via an exponential function model. The tube shape thus defined is used in a volume velocity iterative line analog speech synthesizer, driven by a simulated glottal pulse as described below. The acoustic waveform output of the synthesizer is then analysed with a Markel LPC inverse filter to obtain the magnitude spectrum. Wakita (1972) has described a direct transformation from the area function to Markel inverse filter coefficients which could tremendously shorten the computational task. I have not done this yet, until I can explore more fully



ALASS vowel matching system
Fig. 1

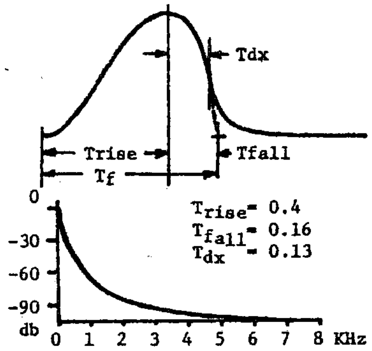


Fig. 2

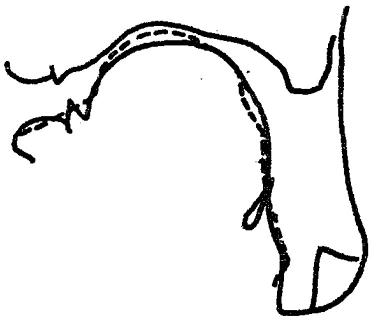


Fig. 3 "heed"

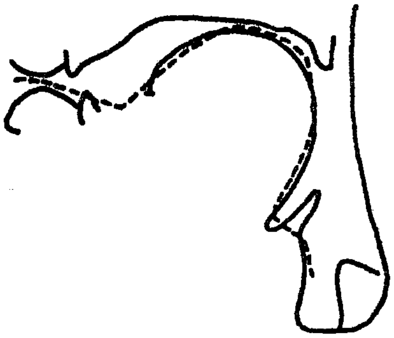


Fig. 4 "who'd"

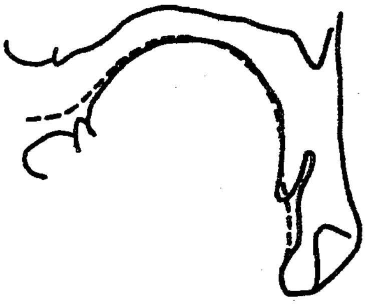


Fig. 5 "head"

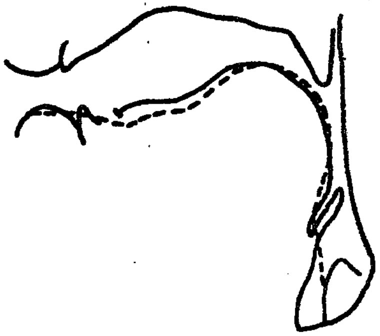


Fig. 6 "hawed"

Solid lines: Tracings from x-rays
Dotted lines: Tongue outline at best ALASS match position

the characteristics of that model regarding losses at the vocal tract walls. The synthesizer presently uses an area dependent loss model which could not be used with Wakita's transformation. As far as the glottal pulse is concerned, however, the spectrum computed via the Wakita transformation could be corrected by adding to it the log spectrum of the chosen glottal pulse shape. Alternatively, this glottal pulse correction could be subtracted once from the target spectrum which could then be directly compared to the Wakita "tube" spectrum. This possibility will be carefully evaluated because of the large possible saving in computational time.

Glottal pulse model

In the line analog synthesizer model, an attempt is made to produce waveform output with as near as possible natural speech quality. Toward this end, a simulated glottal pulse is used to drive the synthesizer based on the curve found to be most natural in listening experiments by Rosenberg (1971). His glottal pulse curve type "B" was modified by marking a point in time on the falling portion and after this time letting the y value decay exponentially toward zero instead of continuing to zero in a sharp corner as in his model. The decay constant is computed such that there is no discontinuity in the slope at time TDX, the decay onset time (see figure 2). As the parameter TDX becomes larger, approaching TFALL, the spectrum slope decreases giving relatively more energy at higher frequencies. With the glottal pulse parameters as shown in figure 2 the glottal volume velocity spectrum has a roll-off of approximately 13 db per octave.

Some vowel matching results

As an initial test of the PARAFAC derived tongue parameter model, the ALASS tracking system was used to recover the articulatory parameter values needed to match the acoustic signals of several of the vowels of the same set originally used in the PARAFAC analysis. The neutral tract shape used was the mean shape over all vowels for the particular speaker, also as used in the PARAFAC analysis. Four figures shown here illustrate the quite high degree of fit in the sagittal plane for the position of best match in the acoustic domain. This degree of fit is even more interesting, considering that a fairly high error remained in the acoustic domain at the best match position. In one case, F_1 was too low by nearly 200 Hz. The discrepancy between a good match in the sagittal plane and a poor match in the acoustic domain can be explained by the fact that the synthesizer contains a fixed mapping of sagittal plane cross-dimension values onto area function values, not tuned to the individual speaker. In fact, such a speaker-tuned mapping could now be constructed by relaxing the tongue constraints in the model and finding the actual area function needed to best match the acoustic signal. However, the fact that the tongue did not move to a position of better acoustic match must be explained by the limitation of the constraints imposed by the parameter model. We see a constraint system operating here which imposes fairly tight restrictions on the tongue motion; in none of the four cases tried could the tongue move away to compensate for a systematic error in another part of the system. On the other hand, if the cross-dimension to area conversion were correct for that speaker, we should expect to see the tongue move to just those positions necessary to attain the target vowels.

References

Harshman, Richard (1970), "Foundations of the parafac procedure: models and conditions for an 'explanatory' multi-modal factor analysis," *Working Papers in Phonetics* 16, UCLA.

Harshman, R., P. Ladefoged, L. Goldstein, and J. DeClerk, (1973), "Factors underlying the articulatory and acoustic structure of vowels," Paper presented to the 86th meeting of the Acoustical Society of America, Nov. 1973.

Markel, John D. (1971), "Formant trajectory estimation from a linear least-squares inverse filter formulation," *Speech Communication Research Labs Monograph*, No. 7, Oct. 71.

Rice, Lloyd (1971), "A new line analog speech synthesizer for the PDP-12," *Working Papers in Phonetics* 17, UCLA.

Rosenberg, Aaron (1971), "Effect of glottal pulse shape on the quality of natural vowels," *JASA* Vol. 49,2, 583-590.

Wakita, Hisashi (1972), "Estimation of the vocal tract shape by optimal inverse filtering and acoustic/articulatory conversion methods," *Speech Communication Research Labs Monograph* No. 9, July 1972.

How May Vowel Systems Differ?

George Papçun

[Paper presented at the 90th Meeting of the Acoustical
Society of America, San Francisco, California,
November, 1975]

How may vowel systems differ? What forces underly them and how can they most meaningfully be described? Various hypotheses have been put forth in the literature to account for vowel systems.

It has been suggested, for example, that within the limits imposed by articulation, that vowels tend to be maximally dispersed in F1 F2 space; that is in the plane of the first formant mapped against the second formant. Liljencrants and Lindblom have explicitly formulated this principle in their article "A Numerical Simulation of Vowel Quality Systems." To check the results of their model, Liljencrants and Lindblom resort to a survey made by Hockett of subjective impressions of language data.

Stevens, in his article "The Quantal Nature of Speech" suggests that there are certain regions in which relatively great articulatory variations produce relatively minor acoustic variations, thus producing acoustic stability in spite of articulatory instability. These are the regions that vowels will be chosen from. Unfortunately, Stevens does not treat real language data in his study.

Quantitative models of vowel systems suffer woefully from a paucity of real language data against which their results may be tested. This is not good. I can even quote you a philosopher who says so: Kant says in the *Critique of Pure Reason*

> "The light dove, cleaving the air in her free flight, and feeling its resistance, might imagine that its flight would be still easier in empty space. It was thus that Plato left the world of the senses, as setting too narrow limits to the understanding, and ventured out beyond it on the wings of the ideas, in the empty space of the pure understanding. He did not observe that with all his efforts he made no advance -- meeting no resistance that might, as it were, serve as a support on which he could take a stand, to which he could apply his powers and so set his understanding in motion."

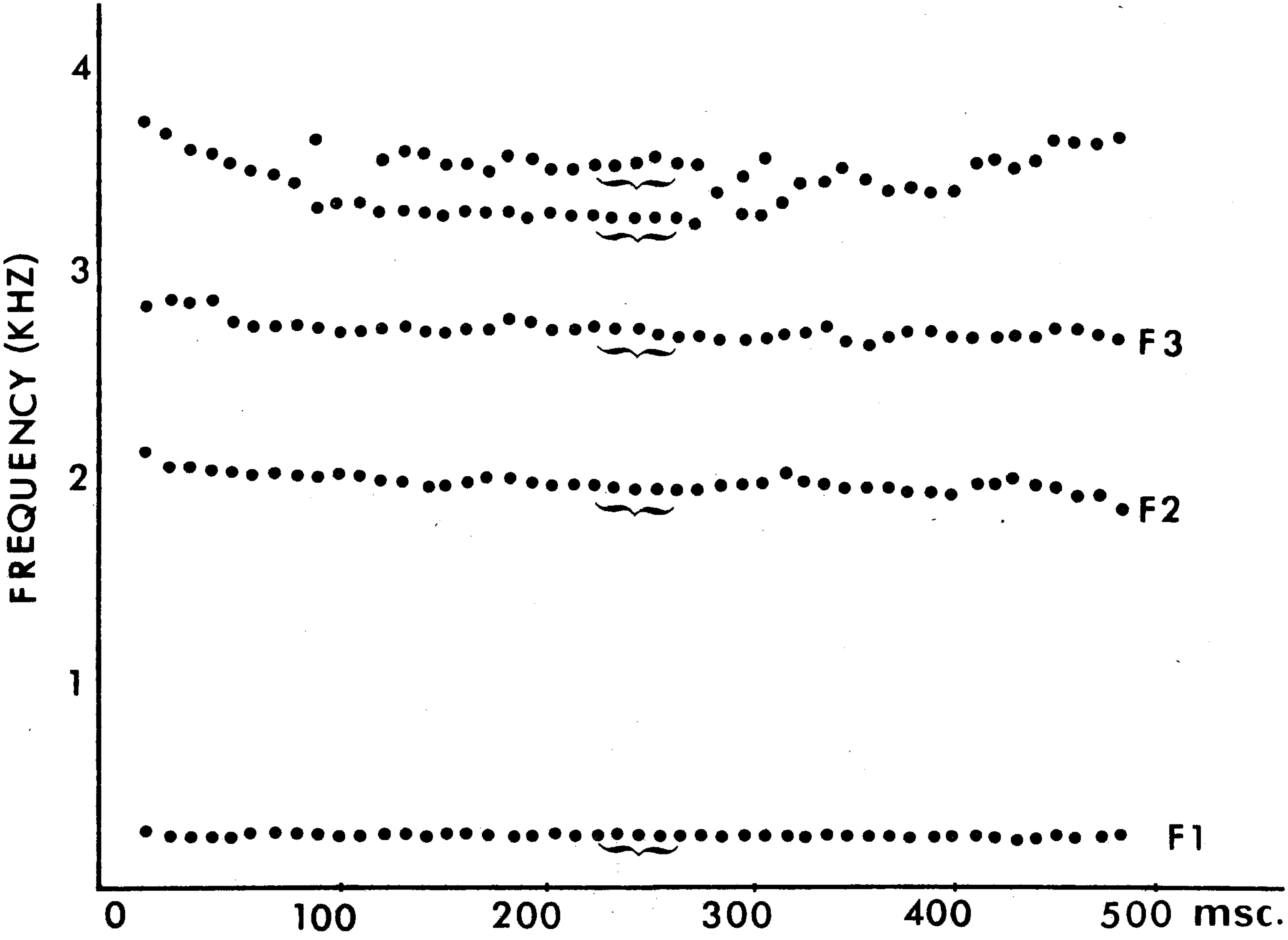


Figure 1. Peaks of the second derivative of successive spectral sections of a vowel. Braces indicate the middle 50 msc. which were averaged to give representative frequency values for each formant. The vowel [i] is shown.

Accordingly, I have set out to test quantitative models of vowel systems by reference to real language data.

In the long run, I hope to assemble rather substantial quantities of data for various languages,but for this report I am presenting only the data for my first 16 subjects: 8 Spanish and 8 Japanese. Spanish speaking subjects are medical students at the Universidad Autonoma de Baja California; Japanese subjects were on a tour of the United States and were taking a short course in English at UCLA, to no effect as far as I could tell. Subjects said the 5 vowels of their languages. Japanese subjects said the names of the appropriate Kanji characters. Spanish subjects said the names of the appropriate letters of the alphabet.

These tokens were LPC spectrum analyzed with Lloyd Rice's WAVES program. The power spectra thus generated at successive 10 msec. intervals were scanned to find positive peaks as reflected in peaks in the second derivative. Figure 1 is a dot plot of the peaks of the second derivative. For each peak the middle 50 msec. of each token as indicated by the curly braces, was averaged for each vowel. These values were taken as representing the formant frequencies. In this study, the first 2 formants only are utilized.

Means were subtracted from the data for each group and the results factor analyzed by Richard Harshman's PARAFAC procedure, a 3 mode factor analysis in which the modes were vowels by formants by persons. At this point an interesting problem came to light: the fit for the Japanese data is a good deal poorer than that for the Spanish data: a correlation .9780 for Japanese as opposed to a correlation of .9975 for Spanish. On looking back over the data, it appears that there was a good deal more instability in the formant measurements for the Japanese than for the Spanish, which may be related to the reaction of the second derivative measurement procedure to bandwidth differences. I know there are subjective differences in voice quality between Japanese and the dialect of Spanish in question, the latter being "sharper," "harsher." Such differences might result from glottal source differences, or differences in pharyngeal stiffness, which might in turn result in such instability.

To derive a normalized estimate of formant frequencies based on best fit in the factor space, the product was taken of loadings on vowels by loadings on formants by mean loadings on persons, and the means previously subtracted out were added back in.

Figure 2 shows the results plotted in an F1 by F2 space with distances made proportionate to the mel scale to depict appropriate perceptions of pitch relations. The circles represent Japanese vowels and the double hatches represent Spanish vowels. Corresponding vowels are connected with dotted lines.

What do we see? First, as any phonetician would have told you, the vowel systems of Japanese and Spanish are not the same. Trivial though this statement is on one level, the significance is that here we have normalized acoustic data which correspond to the phonetician's judgement. Looking further, we see that the /ɛ/ in Japanese is rather lower than the

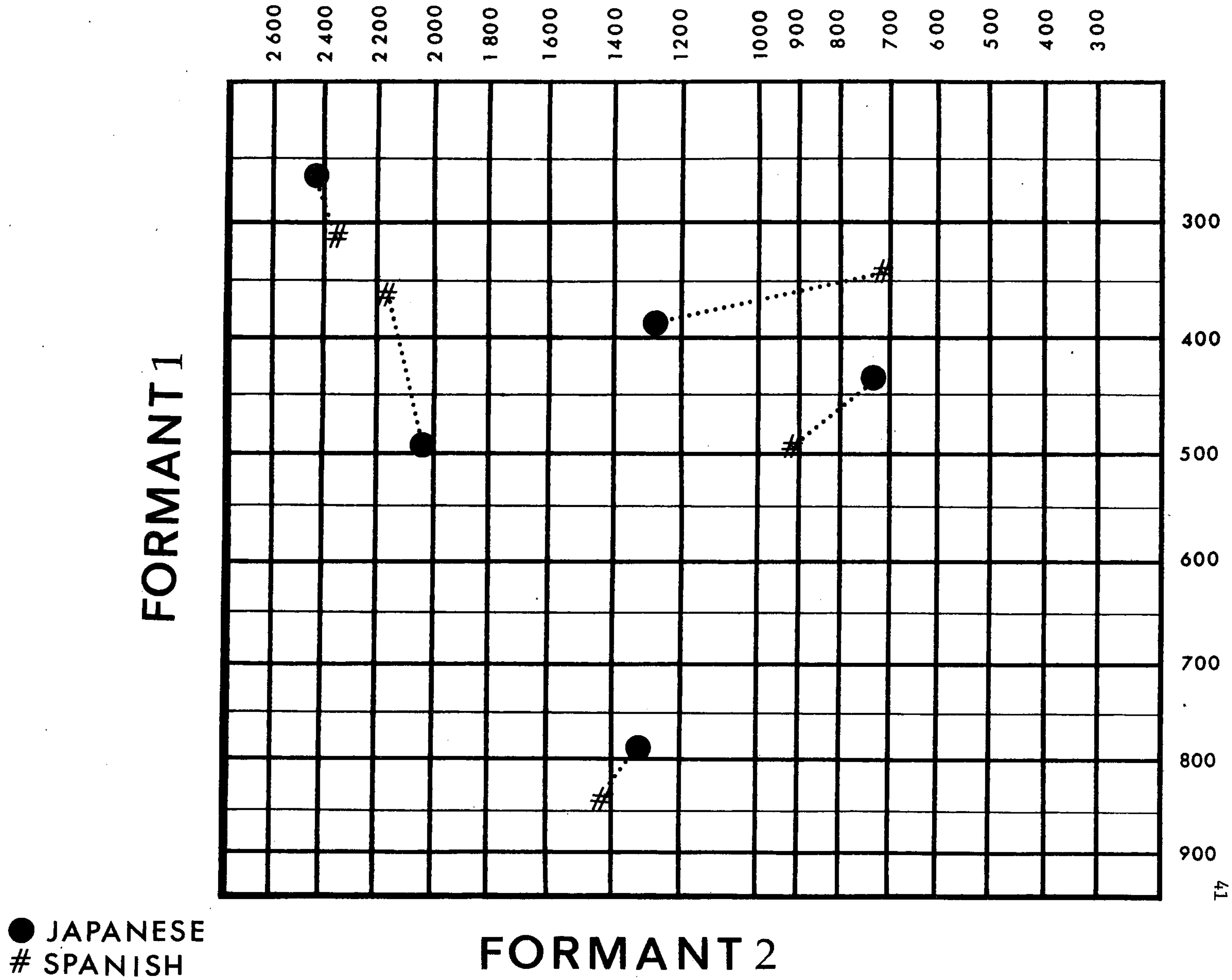


Figure 2. Normalized estimate of formant frequencies based on best fit in factor space. Circles represent Japanese vowels and double hatches represent Spanish vowels. Axes are labeled in Hz. and scaled proportionately to Mels.

/e/ in Spanish. Again, this is as it ought to be by phonetician's judgement. See, for example, Ladefoged's recent text, *A Course in Phonetics*, where just this claim is made.

Note, that the /u/ vowel is further forward for Japanese than for Spanish, presumably because the Japanese version is unrounded. But the region of the high back *rounded* vowel is just one of those places where Stevens would have suggested that vowels should come to rest, congealed in a mass of acoustic stability.

A related sort of problem - both for Stevens and Liljencrants and Lindblom is the extreme proximity of the /i/ and /e/ vowels in Spanish. They are not very dispersed from one another and the /i/ is not even settled into a maximally stable extreme like the Japanese.

However, let us turn now to a plot of the Spanish vowels in factor space. First, I tell you this plot is so pretty - so symmetrical - it has to be right. See the vowel triangle here, leaned a little over, but otherwise impeccable.

It has long been recognized that the distance between the high vowels is too great on an F1 - F2 plot or even an F1 by F2 minus F1 plot. This is reflected for example, in the work of Dale Terbeek and of Jared Bernstein in which it is shown that distance judgements are smaller than would be predicted by an F1 by F2 plot. The same fact comes out in Liljencrantz and Lindblom *Numerical Simulation of Vowel Quality Systems*, in which they find more vowels popping up between /i/ and /u/ than occur in real languages.

Aside from that, we can say generally that all the good points we found in the F1 by F2 comparisons of Spanish and Japanese remain true in the factor plots.

For three reasons, the factor space must be the correct one on which to plot first and second formant data:

1. It is the space which factor analysis shows to be the structure, therefore it must be

2. Two languages give very similar results

3. Phoneticians' judgements are thereby validated.

Future models of vowel systems must be related to such a space.

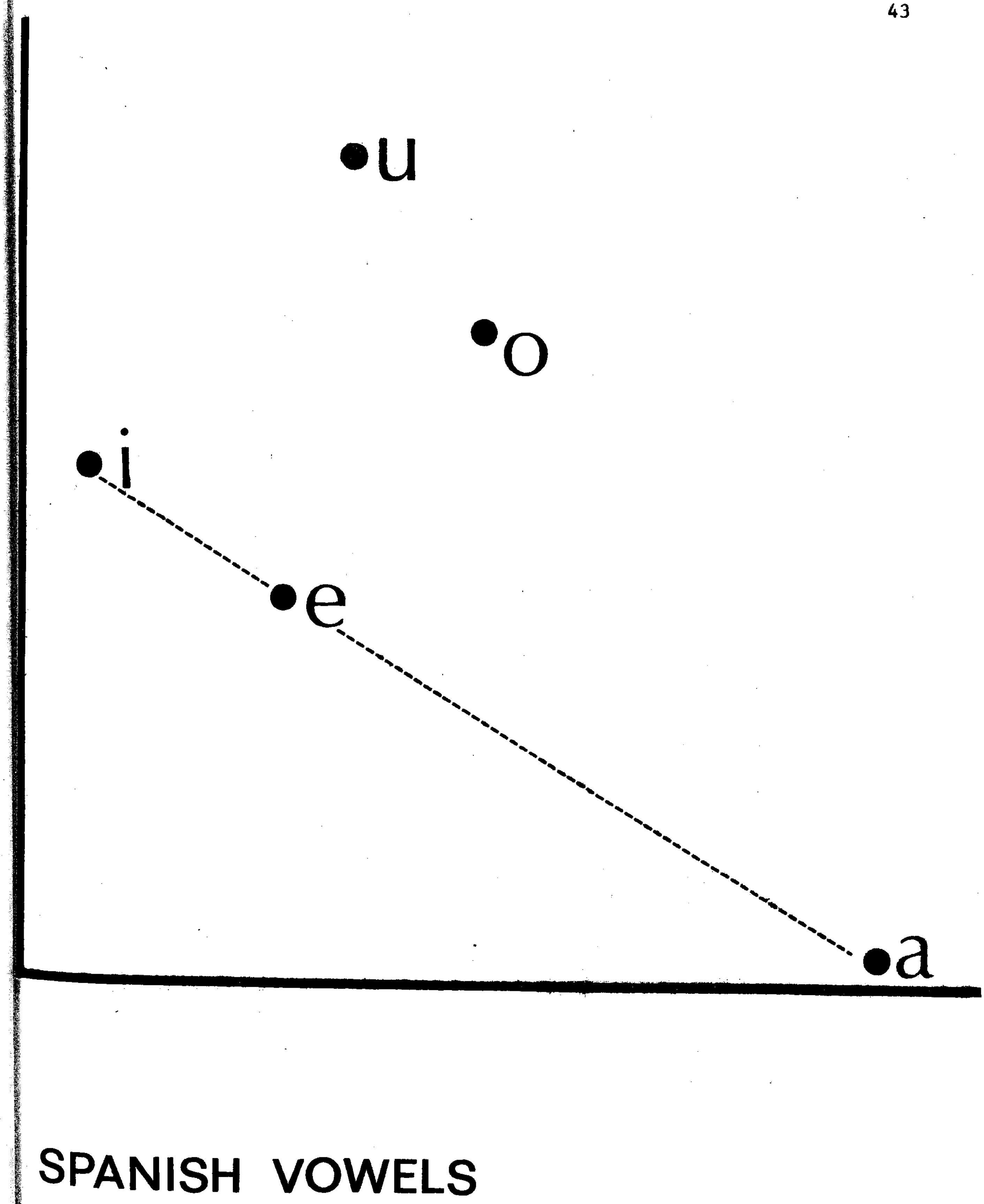


Figure 3. Spanish vowels in factor space.

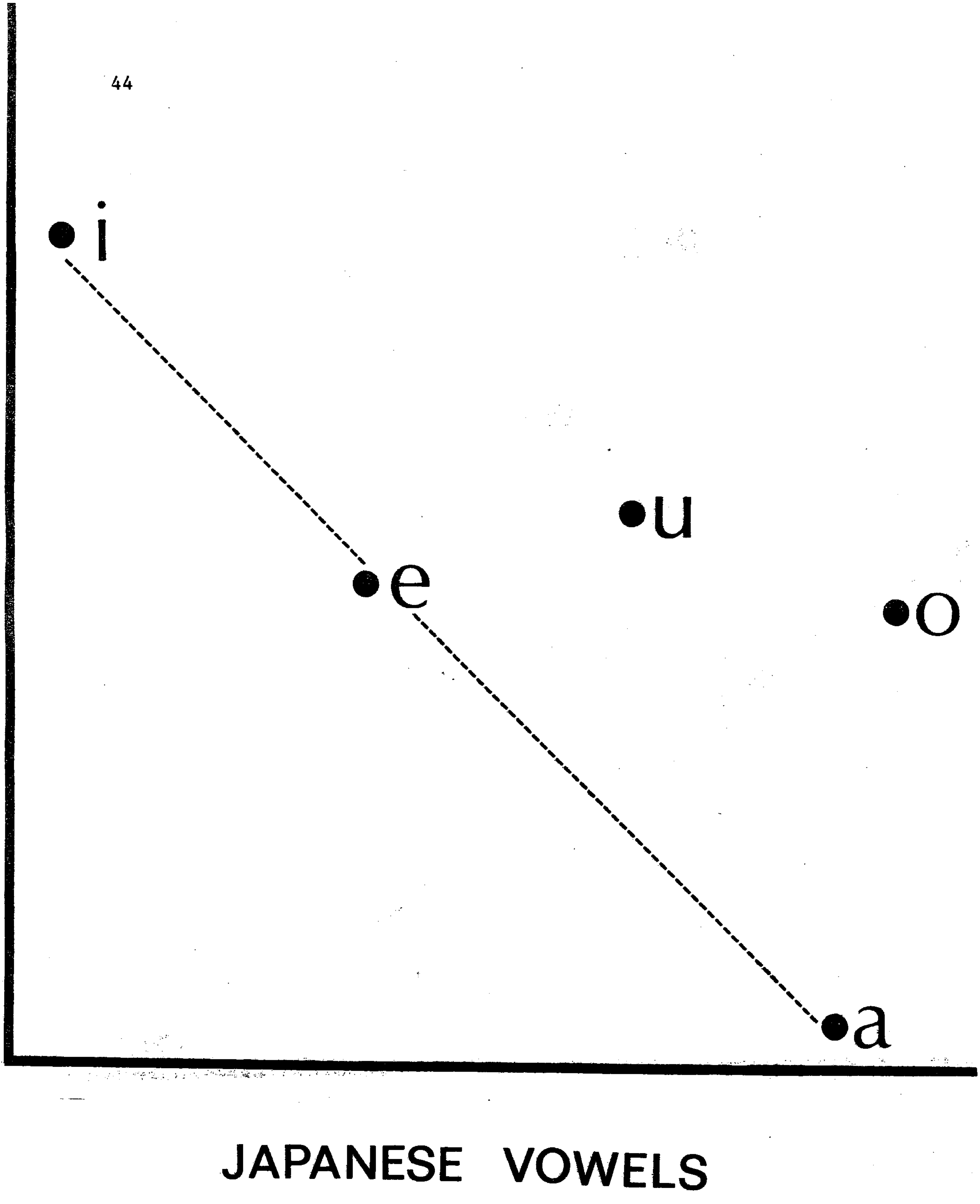


Figure 4. Japanese vowels in factor space.

References

Bernstein, J. (1975) Phonetic distance. Paper presented at Eighth International Congress of Phonetic Science, Leeds, England.

Hockett, C.F. (1955) 'A manual of phonology.' *International Journ. of Amer. Linguistics,* Memoir 11. Baltimore, Md.

Ladefoged, P. (1975) *A Course in Phonetics.* Harcourt Brace Jovanovich, New York.

Liljencrants, J. and B.E.F. Lindblom (1972) 'Numerical simulation of vowel quality systems.' *Language* 48, 839-862.

Kant, I. (1787) *Critique of Pure Reason.* Translated by N.K. Smith, 1958, Modern Library, New York.

Stevens, K.N. (1972) 'The quantal nature of speech: evidence from articulatory-acoustic data. From P.B. Denes and E.E. Davis (Eds.) *Human Communication, A Unified View.* McGraw Hill, New York.

Terbeek, D. and R. Harshman (1971) 'Cross language differences in the perception of natural vowel sounds.' *UCLA Working Papers in Phonetics* 19, Univ. of Calif. Los Angeles.

Vowel Length Before Aspirated Consonants

Ian Maddieson and Jack Gandour

[Paper presented at the 90th Meeting of the Acoustical
Society of America, San Francisco, California,
November, 1975]

It has long been known that, other things being equal, vowels are longer before voiced consonants than before voiceless ones. This phenomenon has generally been considered to be due to an inherent property of the speech production mechanism as it is so widely attested (in German - Meyer 1904; Hungarian - Meyer and Gombocz 1909; French - Gregoire 1911, Belasco 1953, Chen 1970; Italian - Metz 1914; Icelandic - Einarsson 1927; English - House and Fairbanks 1953, Peterson and Lehiste 1960, House 1961; Norwegian - Fintoft 1961; Swedish - Elert 1965; Spanish - Tomas 1916, Zimmerman and Sapon 1958; Dutch - Slis and Cohen 1969; Russian and Korean - Chen 1970; Bengali - Kostic and Das 1972; etc.). A number of different proposals have been made as to what precise mechanism is responsible: some proposals aim only to account for lengthening of vowels before voiced and voiceless consonants (Chomsky and Halle 1968, Chen 1970), while others also aim to account for vowel length differences before consonants of different manners and places of articulation (Belasco 1953, Fischer-Jorgensen 1964, Mohr 1971). We can offer evidence that in Hindi, phonations types other than voicelessness and voicing affect the length of the vowel preceding a consonant. It will be shown that the 'explanations' proposed to account for lengthening before voiced and voiceless stops cannot be extended to account for lengthening before voiceless aspirated and voiced aspirated ('murmured' or 'breathy-voiced') stops.

In Hindi there are four contrasting series of stops: voiceless unaspirated, voiceless aspirated, plain voiced, and voiced aspirated. All four types of stops may appear in word-final position. As shown in (1) below, there are minimal triples; our informant, however, could think of no minimal quadruple except the one in the third column which involves a deleted final vowel in the word for 'obstacle.'

(1)

1.	saat	'seven'	kaat	'cut'	baat	'talk'
2.	saat^h	'companionship'	kaat^h	'wood'	baat^h	'bath'
3.	----		kaad	'embroider'	baad	'later'
4.	saad^h	'balance'	----		baad^h(a)	'obstacle'

To obtain measurements of vowel duration before the four types of final stops, we selected the minimal triple in column 1 and added the nonsense word /saad/. Thus the test items all had initial /s/ and the long vowel /aa/ and final consonants with a dental place of articulation. Although /saad/ is a nonsense word, it is a phonologically possible form.* The other three test items are actually occurring words. A list of the test items (20 tokens of each word) was arranged in random order. Our informant, a native speaker of Standard Delhi Hindi, was recorded reading the test items in the carrier phrase in (2) at normal speech tempo.

(2)

/ab ---- kaho/ 'say ---- now'

The durations of the vowels were measured from a continuous Oscillomink tracing of the waveform from the recording. The results of these measurements of vowel duration are presented in (3).

(3)

Average duration of long vowels (20 tokens)

before /t/	160	msec
before /t^h/	184.5	msec
before /d/	184.75	msec
before /d^h/	196	msec

Vowels are about 25 msec longer before /t^h/ and /d/ than before /t/, 36 msec longer before /d^h/ than before /t/. Vowels before /d^h/ are 11 msec longer than those before /t^h/ or /d/.

The results of grouped-data T tests on these values are given in (4).

(4)

Difference between	Vt	:	Vt^h	$p<.0005$	$t = 6.22$
Difference between	Vt^h	:	Vd	not significant	$t = 0.076$
Difference between	Vd	:	Vd^h	$p<.005$	$t = 3.28$

---

* Dr. Dixit (University of Arizona) pointed out to us that /saad/ is a word meaning "happy" in some non-standard dialects of Hindi.

The difference between the vowels before /t/ and $/t^h/$ is highly significant ($p<.0005$), the difference between the vowels before /d/ and $/d^h/$ is also significant ($p<.005$), but the exceedingly small difference between vowels before $/t^h/$ and /d/ is not significant. We may therefore speak with confidence of three (and only three) different vowel lengths in our data.

It seems likely that vowel length differences of this magnitude result automatically from the differences in the way the following consonants are produced. Given their small size, however, it seems unlikely that they function in Hindi as one of the *major* cues to the identity of the phonation type of the following consonants - which are adequately discriminable on the basis of more obvious cues, for example, the presence of voicing during the initial part of the closure for /d/ and the definite aspirated release of $/t^h/$. In English the difference in vowel length before voiced and voiceless stops is much larger (House and Fairbanks 1953, Zimmerman and Sapon 1958, Peterson and Lehiste 1960, House 1961, Chen 1970, Raphael 1975). In this sense, the differences in vowel length in Hindi are probably not *learned* as they appear to be in English. Our results do, however, suggest that two rules, perhaps 'low-level phonetic rules', operate in Hindi. These may be written as in (5).

(5)
a. vowel adds 1 increment of length before voicing
b. vowel adds 1 increment of length before aspiration

Applying these rules gives the results seen in (6).

(6)
before voiceless unaspirated : 0 increment
voiceless aspirated : 1 increment
voiced unaspirated : 1 increment
voiced aspirated : 2 increments

These findings provide support for the traditional grouping of the voiceless aspirated and voiced aspirated stops together in a class of 'aspirates.' From a diachronic perspective it would seem possible that a 'phonologization' of the rule that lengthens vowels before aspirated consonants could result in long vowels as the historical reflex of aspiration.

While these questions of synchronic and diachronic phonology are of interest, the most interesting aspect of our results is the challenge presented to the various proposed explanations of the cause of intrinsic length of vowels before consonants of different phonation types. Four main theories seem to be represented in the literature on this subject.

One theory (Belasco 1953, Delattre 1962) proposes that the duration of a vowel varies inversely with the force of articulation required for the following consonant. It is assumed that, other things being equal, the 'force' used for each syllable is held constant. As voiceless stops are generally more 'fortis' than voiced ones, this theory predicts that less forceful, i.e. shorter, vowels will occur before voiceless consonants. But it is unlikely that aspirated stops are articulated with less force than unaspirated ones, so this theory fails to predict that vowels are longer before aspirated consonants.

The second theory (Kozhevnikov and Chistovich 1965, Slis 1968, Slis and Cohen 1969, Lehiste 1970) posits a relatively constant duration for the syllable (or some longer processing unit). This duration is maintained by adjusting the duration of segments. In the case of vowels before voiced and voiceless consonants it is the length of the vowel that is adjusted. It is further assumed that voiced consonants are inherently shorter than voiceless ones, and therefore longer vowels appear before the voiced consonants to compensate for their shorter duration. If this theory is to make the correct predictions about our data we should find that the duration of /t/ is longer than the duration of /tʰ/ and /d/, and /dʰ/ is the shortest.

There was a difficulty in measuring the duration of the consonants as we were unable to identify the moment of release for the dental stops in all tokens. Therefore we measured from the end of the preceding vowel to the easily identified release of the /k/ in /kaho/. If we make the reasonable assumption that the duration of the closure for the /k/ segment remained constant, any differing durations measured for the vowel-to-/k/-release interval must be due to the differing phonation types of the dental consonants. The results of these measurements are given in (7).

(7)

Duration of dental stops and /k/

/tk/ : 166 msec

/dk/ : 170 msec

/dʰk/ : 187 msec

/tʰk/ : 192 msec

The plain and aspirated stops are distinguished by a substantial difference in duration. Both the aspirates are markedly longer than the nonaspirates, so that the duration of /tʰk/ is 22 msec longer than the duration of /dk/. Yet the vowels before /d/ and /tʰ/ have the same duration, and the /t/ segment, which has the shortest preceding vowel, is also the consonant with the shortest duration. Therefore a theory of compensatory temporal adjustment cannot explain the pattern of vowel length found in our Hindi data.

The third theory (Halle and Stevens 1967, Chomsky and Halle 1968, Mohr 1971) suggests that vowels before voiced consonants are longer because it requires a complex, and hence *slow* adjustment to move from the glottal configuration for voicing in a vowel (spontaneous voicing) to that for voicing in an obstruent (nonspontaneous voicing, conditioned by a reduced pressure drop across the glottis). The movement from the glottal configuration for voicing in a vowel to the open glottal configuration for a voiceless consonant is a simple, and hence more rapidly executed, adjustment. This theory predicts a long vowel before an aspirated consonant only if the adjustment of the glottis required is a complex one similar to that required in moving to a voiced obstruent. But voiceless aspirated stops are produced with a glottis as 'open' as that required for voiceless unaspirated stops. Thus, the greater length of vowels before aspirates cannot be explained by slow adjustment in the positioning of the vocal folds. Additional problems with this laryngeal adjustment theory have also been pointed out (Chen 1970, Wang 1972).

The fourth theory (Chen 1970) suggests that shorter vowels are due to the wider glottal opening and stronger airflow for voiceless consonants. A more energetic closure is required to overcome the stronger airflow, and this more powerful closure is achieved more rapidly than the closure for the articulators in production of a voiced stop. Chen presents data to show that the closure for a final /p/ in English is indeed carried out in a shorter time than the closure for /b/. If the rate of closure is determined by airflow,and durational differences are the result of faster or slower rates of closure, then the length of vowels before aspirated consonants must also be determined by the airflow rates for these consonants. Thus, as /dh/ has a higher airflow than /d/, this theory predicts incorrectly a shorter vowel before /d^h/ than before /d/, and a vowel before /t^h/ probably shorter than or equal to the duration of a vowel before /t/.

In sum, the Hindi data presented in this paper - showing that vowels are shortest before voiceless unaspirated consonants, are longer but of equal length before voiceless aspirates and plain voiced consonants and longest before the voiced aspirates - shed interesting new light on Hindi phonetics and phonology, furnish fuel for historical speculations about possible origins of lengthened vowels, but, above all, force a rethinking of current explanations of the interaction of phonation type and vowel duration.

References

Belasco, S. (1953) 'The influence of force of articulation of consonants on vowel duration.' *Journ. of the Acoust. Society of America,* 25: 1015-1016.

Chen, M. (1970) 'Vowel length variation as a function of the voicing of the consonant environment.' *Phonetica,* 22.3: 129-159.

Chomsky, N. and M. Halle (1968) *The Sound Pattern of English.* Harper and Row: New York.

Delattre, P.C. (1962) 'Some factors of vowel duration and their cross-linguistic validity.' *Journ. of the Acoust. Society of America, 34,* 1141-1143.

Einarsson, S. (1927) *Beitrage zur Phonetik der Islandischen Sprache.* Oslo.

Elert, C-C. (1965) *Phonologic Studies of Quantity in Swedish.* Skriptor: Uppsala.

Fintoft, K. (1962) 'The duration of some Norwegian speech sounds.' *Phonetica,* 7, 19-39.

Fischer-Jorgensen, E. (1964) 'Sound duration and place of articulation.' *Zeitschrift fur Sprachwissenschaft und Kommunikationsforschung,* 17, 175-207.

Gregoire, A. (1911) 'Influences des consonnes occlusives sur la durée des syllabes precedentes. *Revue de Phonétique,* 1, 260-292.

Halle, M. and K.N. Stevens. (1967) 'On the mechanism of glottal vibration for vowels and consonants.' *MIT Research Laboratory of Electronics Quarterly Progress Report,* 85, 267-270.

House, A.S. (1961) 'On vowel duration in English.' *Journ. of the Acoust. Society of America,* 33, 1174-1177.

House, A.S. and G. Fairbanks. (1953) 'The influence of consonant environment upon the secondary acoustical characteristics of vowels.' *Journ. of the Acoust. Society of America, 25,* 105-113.

Kostic, D. and R.S. Das. (1972) *A Short Outline of Bengali Phonetics.* Statistical Publishing Society: Calcutta.

Kozhevnikov, V.A. and L.A. Chistovich. (1965) *Speech: Articulation and Perception* Moscow-Leningrad . Translated by Joint Publications Research Service: Washington, D.C. No. JPRS 30, 543.

Lehiste, I. (1970) 'Temporal organization of spoken language.' *Working Papers in Linguistics, 4,* 96-114. Ohio State University: Columbus.

Lisker, L. (1974) 'On 'explaining' vowel duration variation.' *Haskins Laboratories Status Report on Speech Research,* 37-38, 225-232. New Haven, Connecticut.

Metz, C. (1914) 'Ein experimentall-phonetischer Beitrag zur Untersuchung der italienischen Konsonanten-gemination.' *Vox,* 23, 201-270.

Meyer, E.A. (1904) Zur Vokaldauer im Deutschen. *Nordiska Studier tillegnade A. Noreen.* 347-356, K.W. Appelbergs Boktryckeri: Uppsala.

Meyer, E.A. and Z. Gombocz. (1909)' Zur Phonetik der ungarischen Sprache.' *Le Monde Oriental* 3, 122-187.

Mohr, B. (1971) 'Intrinsic variations in the speech signal.' *Phonetica* 23, 65-93.

Peterson, G.E. and I. Lehiste. (1960) 'Duration of syllable nuclei in English.' *Journ. of the Acoust. Society of America* 32, 693-703.

Raphael, L. (1975) 'The physiological control of durational differences between vowels preceding voiced and voiceless consonants in English.' *Journ. of Phonetics* 3:1, 25-33.

Slis, I.H. (1968) 'Experiments on consonant duration related to the time structure of isolated words.' *IPO Annual Progress Report* 3, 71-80. Institute for Perception Research, Eindhoven.

Slis, I.H. and A. Cohen. (1969) 'On the complex regulating the voiced-voicless distinction 1.' *Language and Speech* 12, 80-102.

Tomas, N. (1916) 'Cantidad de las vocales acentuadas.' *Revista Filologia Espanola* 3, 387-408.

Wang, W. S-Y. (1972) 'The many uses of Fo.' *Papers in Linguistics and Phonetics to the Memory of Pierre Delattre*, ed. by A. Valdman, 487-503. Mouton: The Hague.

Zimmerman, S.A. and S.M. Sapon. (1958) 'Note on vowel duration seen cross-linguistically.' *Journ. of the Acoust. Society of America* 30, 152-153.

Larynx Height in Kwa

Mona Lindau

[Paper presented at the 90th Meeting of the Acoustical
Society of America, San Francisco, California,
November, 1975]

There is a substantial body of literature on the function of relative larynx height in speech. For a comprehensive review of the literature, see for example Ohala (1973). Variations in the vertical position of the larynx have been related to pitch, to consonant types, and to vowel height. This paper is an attempt to demonstrate another important relationship between larynx height and a feature of vowel harmony in the Kwa languages of West Africa. There is some evidence in data that I have collected that the conclusions reached in this paper also hold true for other language families of West Africa that have similar vowel harmony.

The vowel system of Akan, a language spoken in Ghana, is representative of Kwa languages with this type of vowel harmony. Below are the vowels of Akan.

Set 1			Set 2	
i	u		ị	ụ
e	o		ẹ	ọ
		a		

The distribution of the eight non-low vowels is restricted, so that a word may only contain vowels from one harmony set or the other. On this and the following figures, set 2 vowels are marked with a dot under the letter symbol. The low vowel /a/ is neutral to the vowel harmony and may co-occur with vowels of both sets.

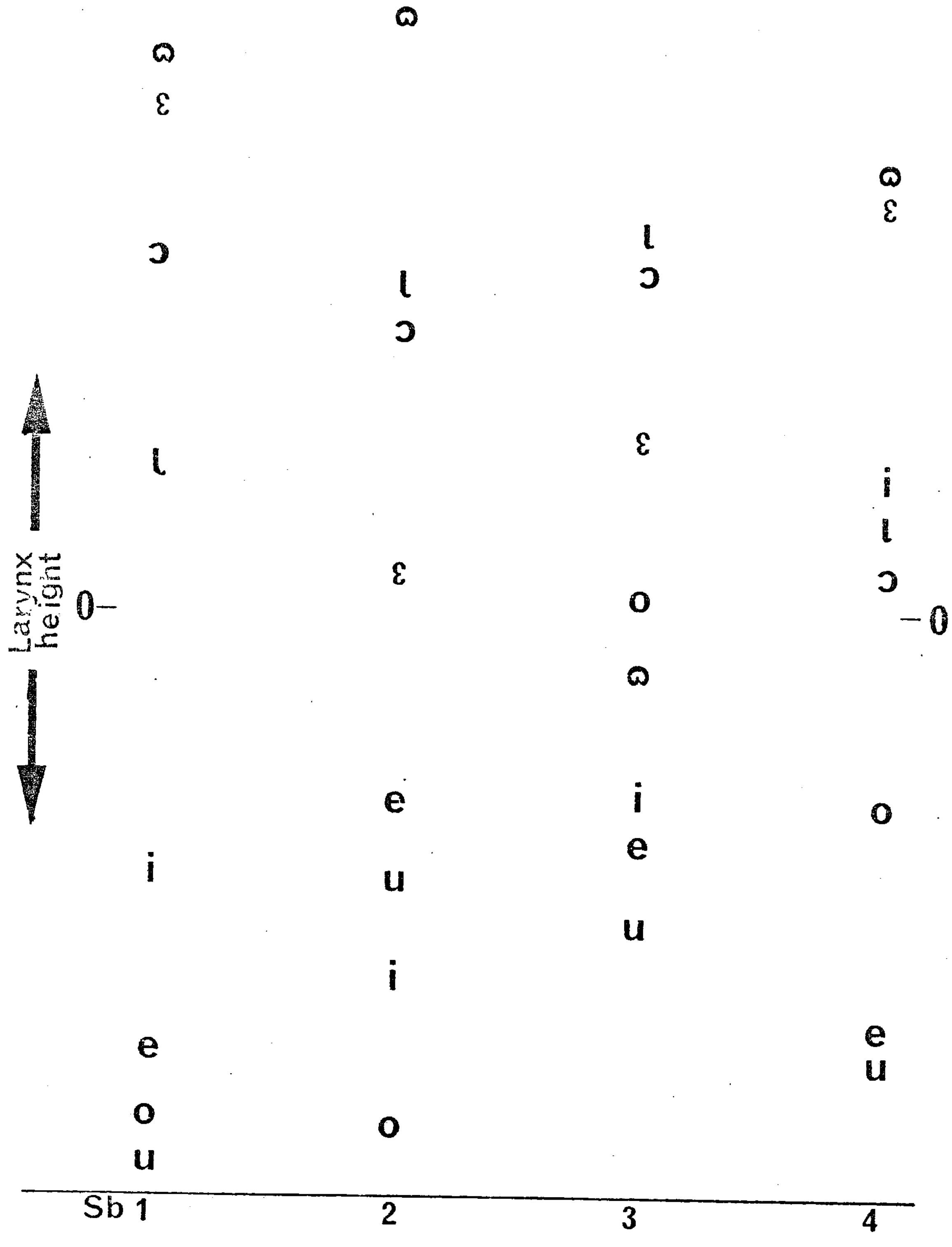


Figure 1. Larynx height of eight vowels in four speakers of Akan. Each vowel point represents the mean of five tokens.

Articulatory properties of these vowels were measured from lateral cine-radiography frames of utterances from four speakers that contained five tokens of each of the eight harmonizing vowels. Larynx height was measured on tracings from x-ray frames as the perpendicular distance from an arbitrary reference line along the floor of the nasal cavity to the superior portion of the thyroid cartilage, which was visible as the bottom of the pharynx on the lateral x-ray frames. Formant frequencies were measured from spectrograms.

Figure 1 is a graph of the mean larynx height of the vowels of each speaker. It is quite evident that a major function of the larynx in this language is to contribute to keeping the two vowel sets apart. The larynx is higher for all set 2 vowels than for all set 1 vowels in all subjects, except for two minor discrepancies. In subject 3, the positions of /o/ and /ụ/ are reversed, but every set 2 vowel is still much higher than its set 1 counterpart. In subject 4 the vertical positions of /i/ and /ị/ are much the same. But the larynx height of the other vowels clearly separate the two sets.

The mechanism for separating vowels in different sets involves more than the height of the larynx. The articulatory positions of the four vowel pairs are illustrated by superimposed tracings of these vowels in a speaker of Akan in figure 2. Front and back vowels are superimposed separately. Set 1 vowels are drawn with an unbroken line, set 2 vowels with a dashed line. Comparing vowels of different sets we find that over and above the differences in larynx height, there are differences in the position of the root of the tongue. (There are also differences in the height of the tongue, but statistical analysis shows that these are predictable from the position of the tongue root, and not significantly different for the two sets.)

Analysis of variance was used to determine the significance of the differences between the two vowel sets as to both parameters of pharyngal width and larynx height. For both parameters there are highly significant differences ($p < 0.001$) between the two sets. Results from regression analysis of the relationship between pharynx width and larynx height show that the two parameters correlate very well, with values of Pearsons r ranging from 0.90 to 0.55 for the four subjects. The correlation is negative: as the width of the pharynx increases, the height of the larynx decreases, that is, as the tongue root moves forward, the larynx moves downwards.

In another data set from an East African language, Ateso, with a structurally similar vowel harmony, there is no vertical displacement of the larynx accompanying the variation in pharynx width, so larynx height does not always accompany pharynx width. Moreover, there are no intrinsic connections between the larynx and the tongue root that would necessitate the larynx moving vertically down as the tongue root moves horizontally forwards. Thus the height of the larynx is controlled independently of the tongue root.

By combining the advanced tongue root with a lower larynx position, the speaker accomplishes a larger variation of the size of the pharyngal

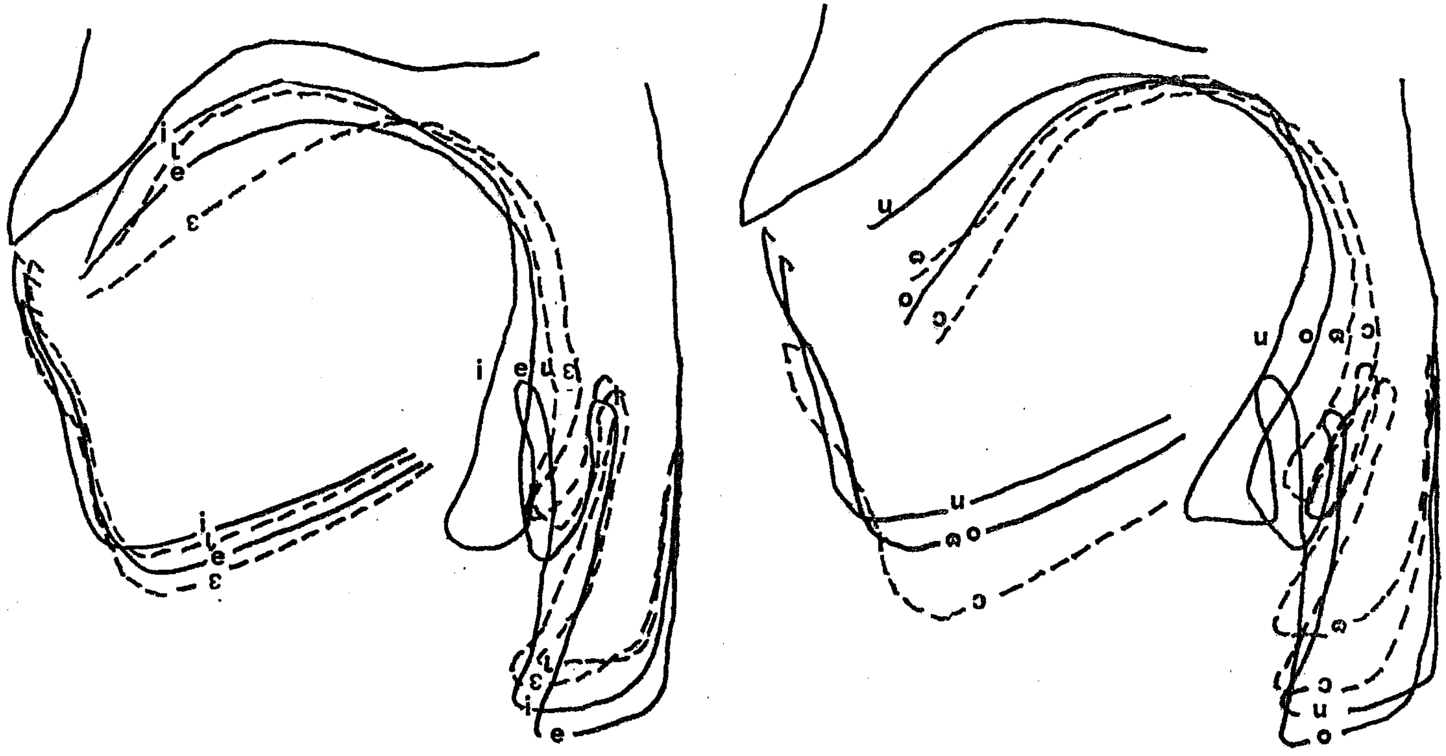


Figure 2. Superimposed tracings of front and back vowels of a speaker of Akan.

cavity, than if he varied just the tongue root position alone. One would thus expect larger acoustic effects than if there were no changes in the position of the height of the larynx. And indeed, in Akan, and in other Kwa languages that I have investigated, the function of the variations in larynx height is to help increase the acoustic distance between different sets of vowels.

A comparison of the observed formant frequencies of the two vowel harmony sets in Akan shows that varying the size of the pharyngal cavity affects both the first and the second formants. The effects on the third formant were negligible. Figure 3 is the usual kind of plot of the mean frequencies in mel of the first formant against the second formant of each of the four speakers. Each formant frequency on this chart represents an average of the five tokens of each vowel. The first formant is generally considerably higher for set 2 vowels than for their set 1 counterparts, that is, the set 2 vowels are lower on the chart. The second formant is lower for set 2 vowels than for their set 1 counterparts among the front vowels, but it does not differ much for different harmony sets among the back vowels. When the differences in each formant between set 1 and set 2 vowel pairs are calculated in terms of percentage differences, one finds that the relative differences are much larger for the first formant than for the second formant. Changing the size of the pharyngal cavity thus changes the first formant more than the second formant, that is, the vertical position on the vowel on a vowel chart is more affected than its position on the horizontal dimension.

An attempt was made to tease out the relative contribution of larynx height and pharyngal width to the changes in the formant frequencies. Vowels were synthesized using a line analogue speech synthesizer. This vocal tract model has been described by Rice at previous meetings of the Acoustical Society. This model divides the vocal tract into 18 sections and calculates the formant frequencies for a given vocal tract shape. It is possible to change the length of the pharynx by deleting the first section above the larynx. This is equivalent to raising the larynx 1 cm.

An [i]-shape and an [u]-shape were generated on the model, and the tongue root was systematically varied from an advanced position to a retracted position in several steps under two conditions (as illustrated in figure 4). Condition 1 was with the larynx in "normal", or low, position, and condition 2 was with the larynx raised 1 cm. The effects on the formants were found to be similar to those observed in the Akan data. Raising the larynx makes virtually no difference to either the second or the third formant. The important contribution of variation in larynx height is to the frequency of the first formant. Figure 5 is a graph of the effect of tongue root movement on the first formant in an [i]-shape and an [u]-shape, under the condition of a low larynx (unbroken lines) and under the condition of a raised larynx (dotted lines). The frequency of the first formant is more affected when the tongue root is retracted than when it is advanced. Secondly, the effects from enlarging the pharyngal cavity are larger for [i] than for [u]. The

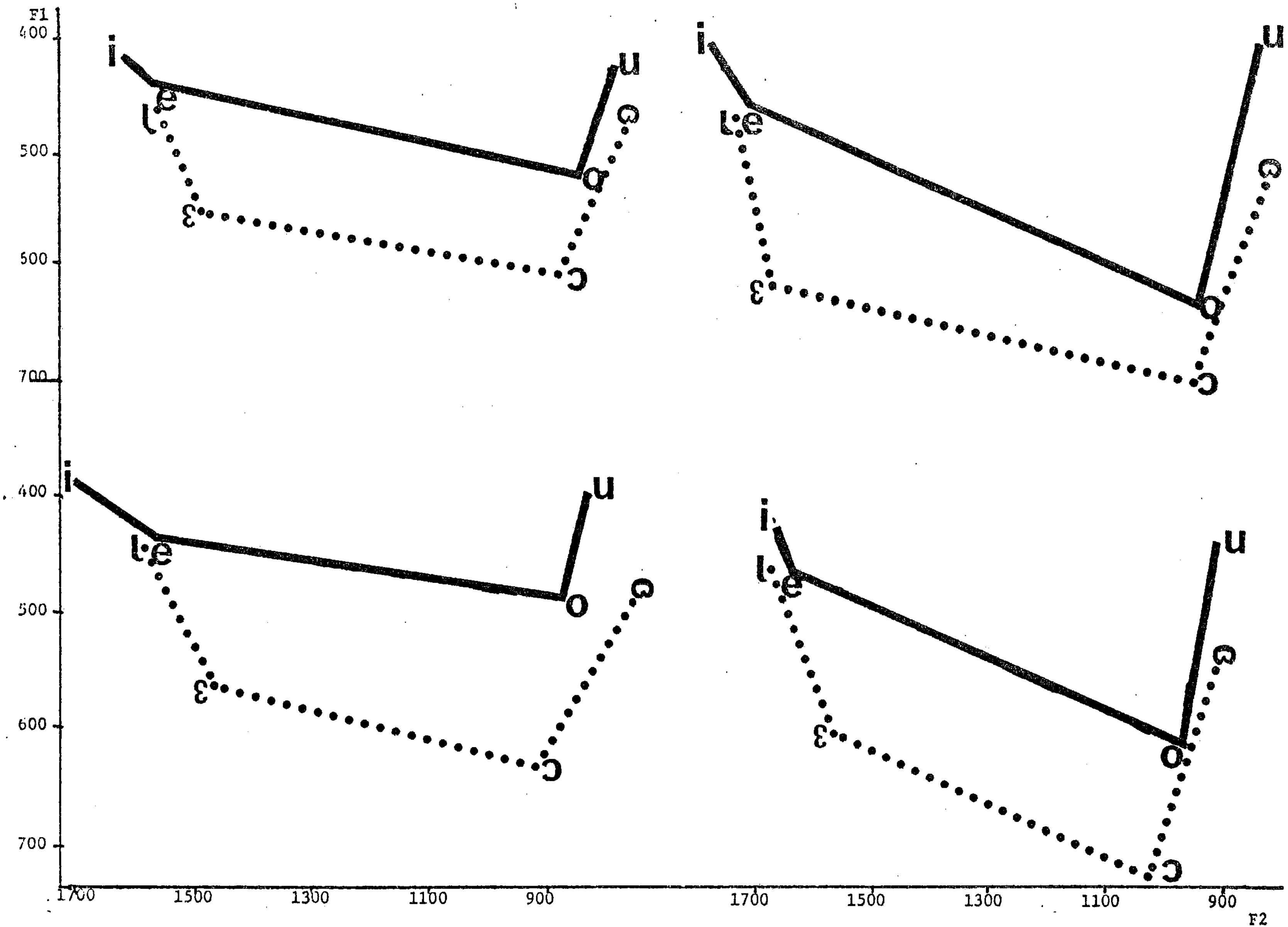


Figure 3. Formant frequencies in mel of the second formant and the first formant of four speakers of Akan. Each point represents an average of five tokens.

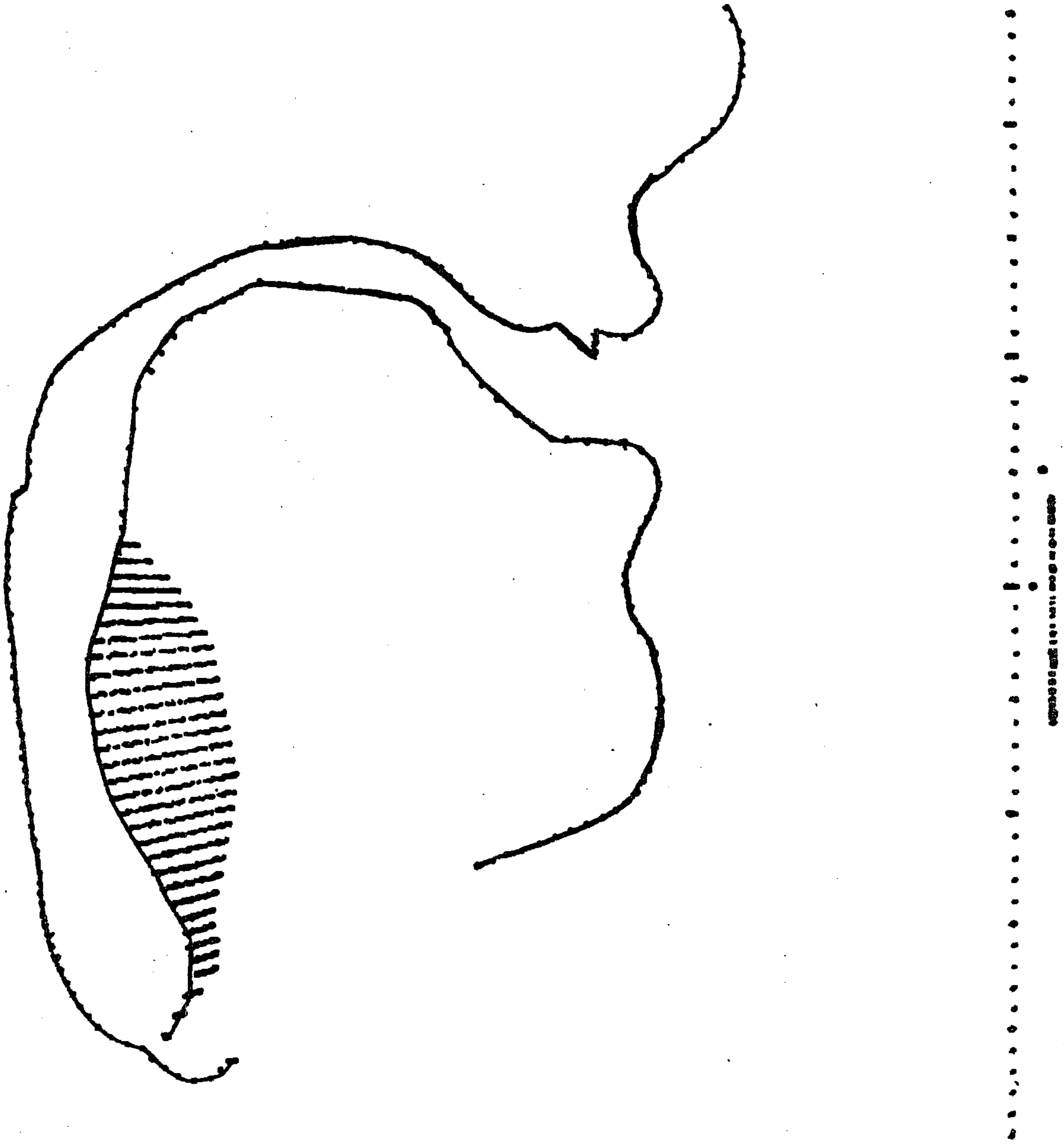


Figure 4. An [i] shape generated on the LASS model of the vocal tract with variable larynx height.

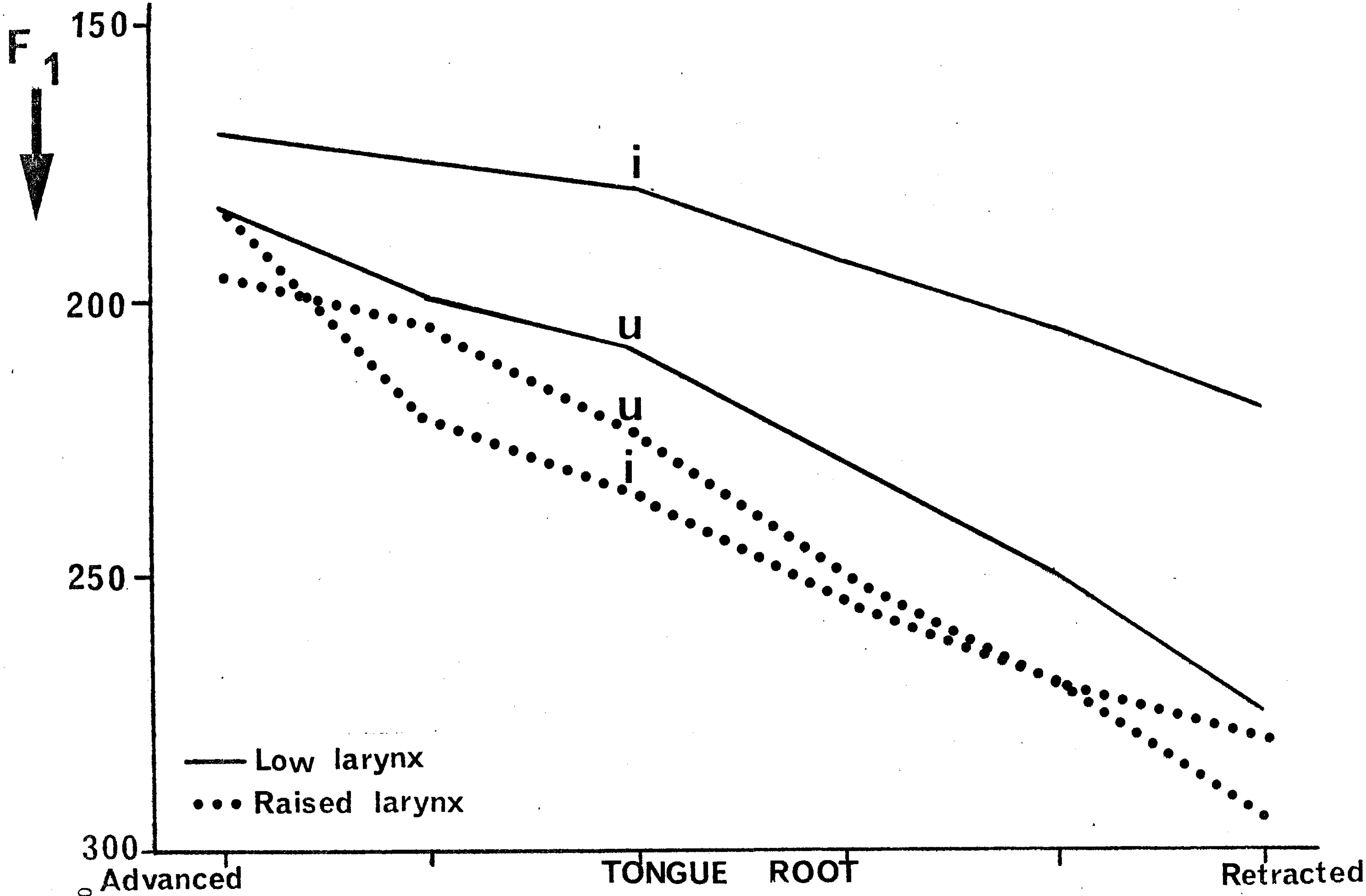


Figure 5. Graph showing the effect on F_1 of varying the tongue root position under conditions of lowered and raised larynx in an [i] and an [u].

largest effects for both vowels are accomplished by going from an advanced tongue root with a low larynx to the retracted tongue root with a raised larynx -- exactly as in fact happens in Akan and other Kwa languages. Changing the height of the larynx here has a definite function in making the vowels of different harmony sets more different from each other acoustically than if the larynx height were kept constant.

This appears to be the major function of larynx height in Akan. The height of the larynx is not significantly different in any systematic way between high and non-high vowels within each set, nor between front and back vowels.

The relationship between pitch and larynx height was also studied for two male speakers of Akan. Akan has high and low tones, but all the investigated vowels were spoken on a high tone. The fundamental frequencies of the different high tone vowels varied within a range of 25 hz. There was, however, no significant correlation between larynx height and pitch for set 1 vowels. For set 2 vowels there was a weak tendency for the fundamental frequency to increase as the larynx height decreased (pearsons r = -0.12 for one speaker and -0.40 for the other speaker). In general, pitch and larynx height do not correlate within one tonal category.

In a study of larynx height in Thai, Gandour and Maddieson (1975) found that tonal category accounted for some of the variation in larynx height. So before dismissing any possible relationship between larynx height and tone in Akan, a study would have to be done to include low tone vowels as well.

To summarize: The larynx height of vowels in Akan, and presumably other Kwa languages cannot be correlated with pitch, nor with vowel features of Height and Backness, but can only be correlated with the vowel harmony feature.

References:

Gandour, J. and I Maddieson (1975) 'Measuring larynx movement in Standard Thai using a cricothyrometer' Unpublished paper, Linguistics Department, University of California, Los Angeles.

Lindau, M. (1975) 'Features for Vowels' *UCLA Working Papers in Phonetics 30*

Ohala, J. (1973) 'Explanations for the intrinsic pitch of vowels' Monthly Internal Memorandum (January) 9-26), Phonology Laboratory, University of California, Berkeley.

Acknowledgments:

I wish to thank Peter Ladefoged for many helpful comments and suggestions, and Jack Gandour, Ian Maddieson, and Bill Ewan for many helpful rapping sessions on the subject of larynx height. I also wish to thank Leon Jacobson for contributing his x-ray data on Ateso.

The Natural Mnemopath: or, What You Know About Words You Forget

Catherine Browman

(Paper presented at the 86th meeting of the
Acoustical Society of America)

What are the parameters available for use in retrieving words from memory? Linguistic errors can be useful sources of data for this question. A particularly useful data source is the tip-of-the-tongue phenomenon (TOT), in which the individual is trying to recall a word, but can't quite remember it. In 1966 Brown and McNeill studied the TOT phenomenon under experimental conditions in which the forgetting was induced. In order to eliminate any artificialities introduced by the experimental situation, the present study concentrates on naturally occurring TOT phenomena, especially those in which an incorrect *approximation* word is produced for the desired *target* word. The ways in which each approximation is similar to its target reveal what is known about the target before it is completely recalled, and thus suggest some parameters of storage.

During 1973, the author collected 113 target-approximation pairs from 24 different individuals, with the largest number of pairs from any single individual being 21. The target and approximation word from each pair were compared to determine what they had in common. (Examples of data pairs are shown in Table 1.) In order to determine the significance of the common elements in each data pair, it was necessary to establish the chance level - that is, before stating that a particular similarity between the approximation and the target revealed some knowledge about the target, it was necessary to find how likely that similarity was to occur by chance. Chance level was determined by establishing the similarities between the two members of each of 113 random word pairs, where the 113 random word pairs came from the same population as the original word pairs. Specifically, the 113 random word pairs were generated by randomly reassigning each approximation to a different target. Thus for both the random and observed data pairs, the same 113 words served as targets and another 113 words served as approximations - just the pairings differed. The chance level determined by the amount in common between approximation and target of the random word pairs was used to test the significance of similarities between members of the observed word pairs throughout the study.

		OBSERVED DATA	RANDOMIZED DATA
	TARGET	APPROXIMATION	APPROXIMATION
DATA EXCERPTS	JENNIFER	JESSICA	COINCIDENCES
	CROQUET	CRICKET	WATERWEEDS
	HARTSHORN	ANTLER	DIFFERENCES
	POMPEII	PTOLEMY	DIGRAM
	DISINTEGRATION	DEGRADATION	MEYERS
	IMPRECATE	RINALDI	BABBITT

Table 1. Examples of target words, observed approximations, and randomly assigned pairings.

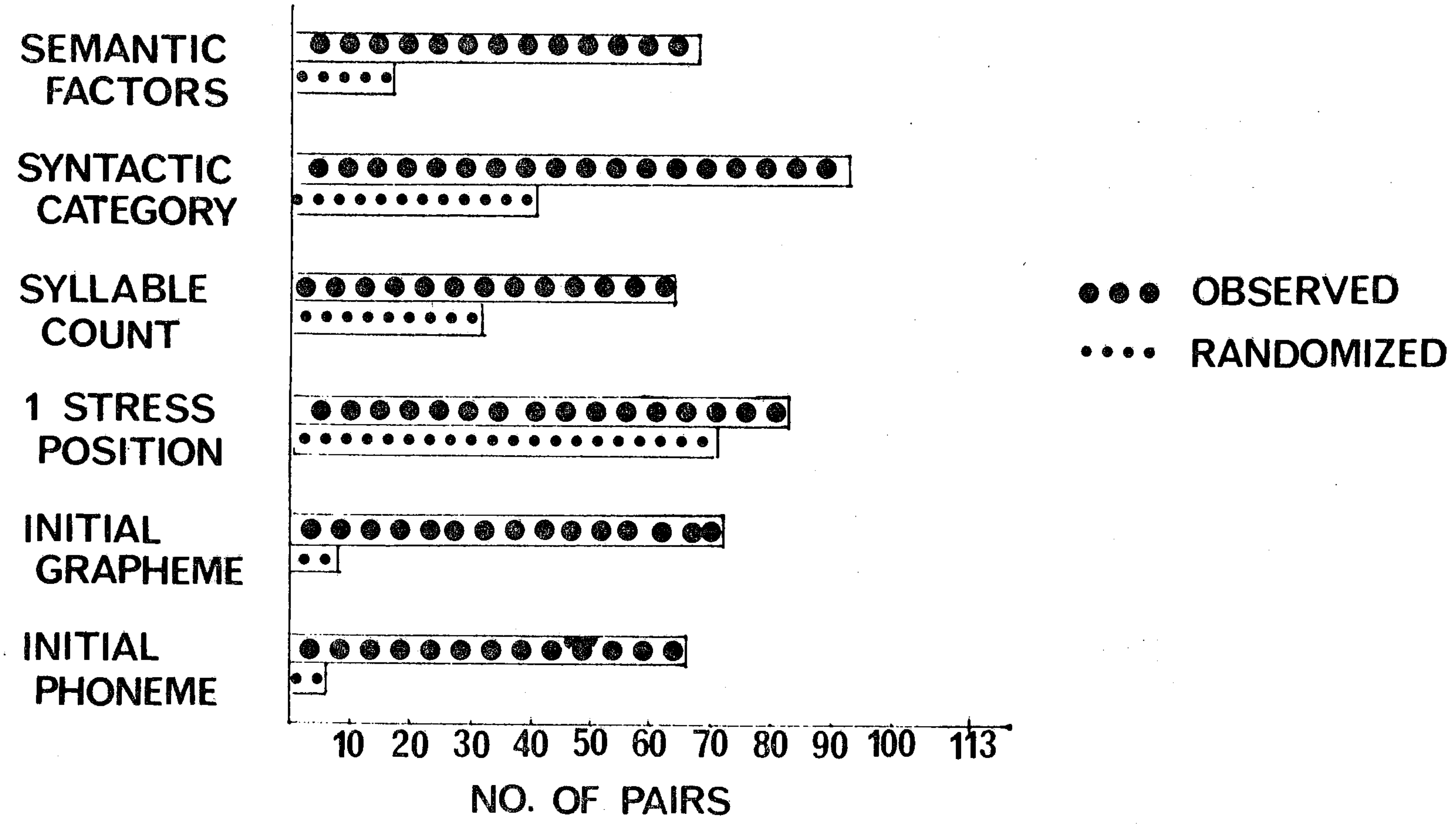


Figure 1. Identity between target and approximation.

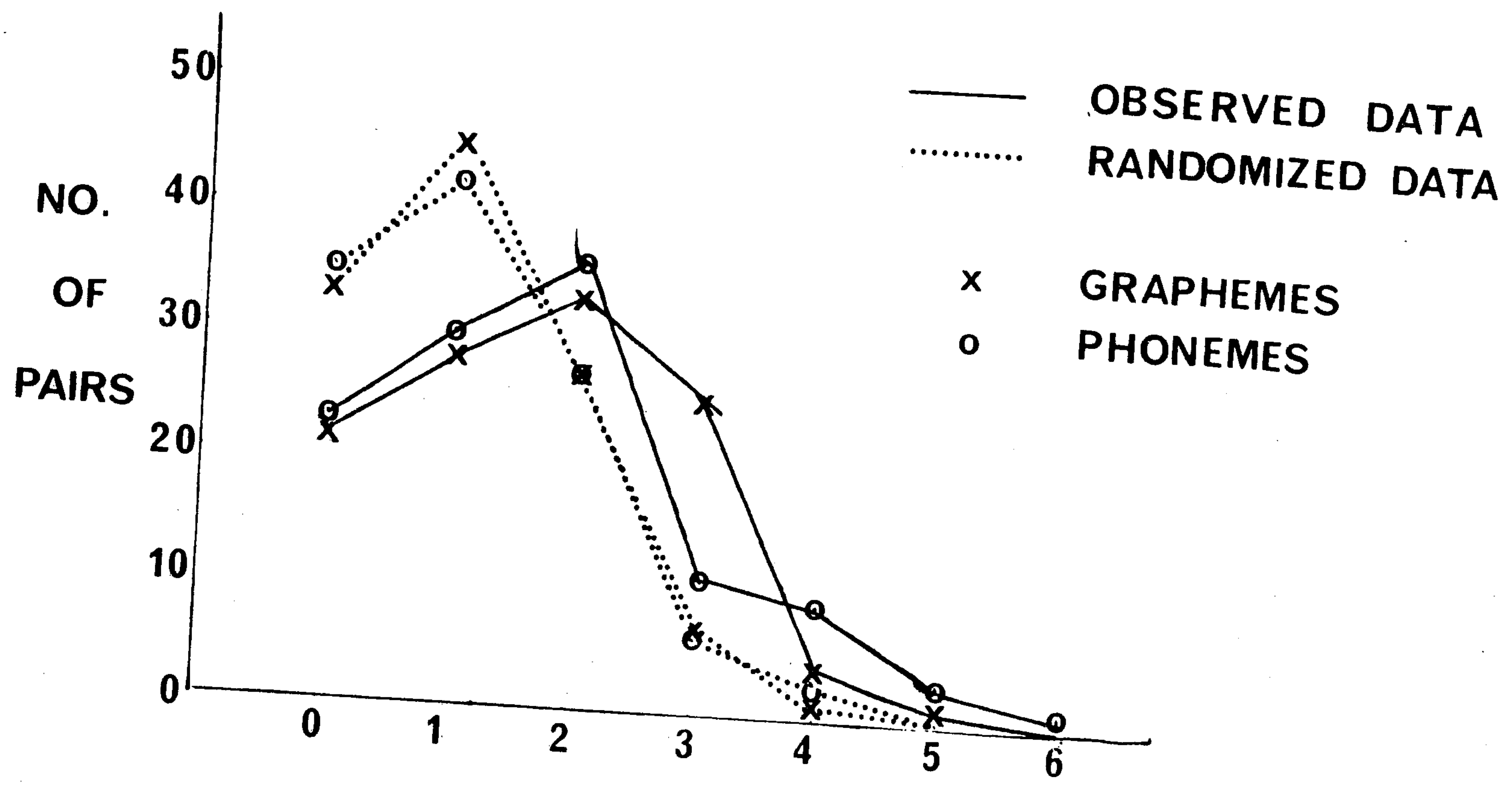


Figure 2. Number of units in common between target and approximation.

The amount in common between target and approximation of the observed word pairs was determined for seven factors: semantic factors, syntactic category, number of syllables, position of primary stress, initial grapheme, initial phoneme, and word initial consonants, both graphemic and phonemic. For all seven factors, the observed data pairs had more in common than the randomized pairs - the observed randomized difference was significant at $p<.01$ for every factor except position of primary stress. The results are summarized in Figures 1 and 2. 68 of the 113 observed data pairs were semantically similar (as compared to 15 of the 113 randomized pairs) where the speaker's judgment of similarity was generally used, occasionally supplemented by the author's judgment. For example, *Pompeii* and *Ptolemy* were judged as semantically similar (both ancient, classical) while *imprecate* and *Rinaldi* were judged as semantically dissimilar. 93 of the observed data pairs belonged to the same syntactic category (compared to 41 randomized pairs). Of these 93 pairs, 81 were noun-noun pairs (compared to 40 randomized noun-noun pairs). 64 of the observed data pairs had the same number of syllables (compared to 32 randomized pairs). 83 of the observed data pairs had the primary stress in the same position (compared to 71 of the randomized pairs) where the identity of position was determined by either one of two conditions being true: the primary stress of both target and approximation were on the same syllable starting the count either 1) from the last syllable of both or 2) from the first syllable of both. Thus both the pair *Jennifer* - *Jessica* and the pair *disintegration* - *degradation* were categorized as having primary stress in the same position in both target and approximation. The accuracy of position of primary stress was independent of the accuracy of the number of syllables - that is, the condition of correctly approximating the number of syllables in the target word did not increase the likelihood of correctly approximating the position of primary stress, and vice versa. 72 of the initial graphemes and 66 of the initial phonemes were correctly approximated in the observed data pairs (compared to 8 and 8 for the randomized pairs). For the observed word pairs, the mean number of non-initial consonants (both graphemic and phonemic) in common was 1.7, compared to 1.1 for the randomized pairs. The difference in the distribution of non-initial consonants in common between the observed and the randomized data was significant with $p<.01$.

In summary, the present data indicate a number of factors are known about a desired target word prior to complete recall: semantic factors, syntactic category, number of syllables, and what the consonants are - both initially and non-initially, graphemically and phonemically. Apparently the position of the primary stress is not, on the whole, known prior to recall. Finally, there is a slight tendency for the graphemic realization of a consonant to be recalled more often than the phonemic realization.

Acknowledgments

Many thanks to Mike Goodman of Bell Labs and Richard Harshman of UCLA for statistical advice. This work was supported by Grant USPHS NS 9780.

References

Bolinger, Dwight (1961), "Verbal evocation," *Lingua* X:2, 113-127.

Brown, Roger, and David McNeill (1966), "The 'tip of the tongue' phenomenon," *Journal of Verbal Learning and Verbal Behavior* 5, 325-337.

Freedman, J.L. and T.K. Landauer (1966), "Retrieval of long-term memory," *Psychon. Sci.* 4:8.

Yarmey, A. Daniel (1973), "I can recognize your face but I can't remember your name: Further evidence on the tip-of-the-tongue phenomenon," *Memory and Cognition* 1:3, 287-290.

Frigidity or feature detectors-slips of the ear

Catherine P. Browman

[Paper presented at the 90th Meeting of the Acoustical
Society of America, San Francisco, California,
November, 1975]

When people think of speech errors they usually think of production errors,or slips of the tongue spoonerisms such as saying "queer old dean" instead of "dear old queen." Other kinds of speech errors, however, also occur - for example, perceptual errors, or slips of the ear.

Over the last year, I have collected, with substantial aid from friends and acquaintances, about 150 slips of the ear. Ninety-three of these have been analyzed, with special emphasis on lexical structure such as word boundaries. In Figure 1, examples are shown of the kinds of change in lexical structure that occur in slips of the ear.

Of the 162 words in the analyzed corpus, by far the majority show no change in the placement of the word boundary between the spoken and the perceived utterances. Thus, in example 1) *bard*, in the phrase "wise old bard", was perceived as *fart*, as in "wise old fart." Four kinds of change in lexical structure do occur, however. These changes are indicated in the examples by the circled words boundary or syllable number, where the numbers refer to the syllables that correspond in the spoken and perceived utterances. Thus in example 2, the final word boundary of the first word shifted from between syllables 3 and 4, in the spoken utterance *notary public*, to between syllables 2 and 3 in the perceived utterance *Nota Republic*. Example 3 shows one of the 14 utterances of word boundary insertion. In this case the spoken utterance *herpes zoister* was split up into 3 words by the insertion of a word boundary between syllables 1 and 2. There were 12 instances of deletion of the final word boundary, as in example 4. Here, the final word boundary between syllables 2 and 3 in the spoken utterance *popping really slow* was deleted in the perceived utterance *prodigal son*. Note also that syllable 4 was deleted - the perceived utterance has only 4 syllables instead of 5. Finally, there were 8 cases of syllable insertion, as in example 5. Here a syllable was inserted before the first syllable of the spoken utterance *Freudian* to get, with segmental changes, *accordian*.

These perceptual errors involving lexical structure change are particularly interesting insofar as they relate to higher level perceptual processes. Speech perception can perhaps best be thought of as a process of lexical retrieval which is only partially mediated by the acoustic

TYPES OF CHANGE IN LEXICAL STRUCTURE IN PERCEPTION OF SPOKEN WORDS:

FINAL WORD BOUNDARIES

1) SAME

bard → fart
1# 1#

2) SHIFT

notary public → nota republic
1 2 3(#)4 5# 1 2(#)3 4 5#

3) INSERTION

herpes zoister → her peas oyster
1 2 # 3 4 # 1(#) 2 # 3 4 #

4) DELETION

popping really slow → prodigal son
1 2(#)3 4#5 # 1 2 3#5 #

SYLLABLES

5) INSERTION

freudian → accordian
2 34 # ① 2 34 #

Figure 1: Slips of the Ear-types of Lexical Structure Changes.

signal. Other higher predictive levels also contribute to the process of lexical retrieval, such as prior syntactic analysis, frequency, and context. Segmental errors in perception, then, may be caused both by low-level acoustic errors and by the retrieval of the wrong lexical item as influenced by higher-level processes. The less the information in the acoustic signal is used in determining the segments, the less closely the perceived segments will correspond to the spoken segments. Thus those slips of the ear that show the clearest evidence of higher level processes should also have the most segmental errors.

Slips of the ear involving changes in lexical structure are the likeliest candidates for maximal higher-level involvement. Errors involving syllable insertion are the clearest case, since *no* information in the acoustic signal is present for the inserted syllable. Thus, words that are perceived with an inserted syllable should also be perceived with many segmental errors. Next in line should be words with errors involving boundary deletion, since here the word retrieved is longer than the word uttered, thus indicating the possibility of inattention to the acoustic signal for a longer time. Final word boundary shifts to the right should be about on a par with word boundary deletions. The least effect should be shown by words with errors involving word boundary insertions and final boundary shifts to the left, since in these cases the word retrieved is shorter than the word uttered, and thus the duration of possible inattention is less. Thus, to illustrate, compare examples 3 and 4 in Figure 1. In number 3, the spoken word *herpes* has been perceived with a word boundary inserted, resulting in the shorter lexical items *her* and *peas*, both segmentally identical to the spoken utterance. Compare this to example 4, where the word boundary following *popping*, has been deleted. The perceived word *prodigal*, which is longer than the spoken word *popping* also has fewer segments in common with *popping*.

To test whether in general these candidates for higher level - introduced errors showed more segmental errors, I correlated the percent of syllables showing errors with the type of lexical structure change. Figure 2 shows the results. Since only 3 words showed word boundary shifts, I did not include their results. Comparing spoken words which were perceived with inserted word boundaries, to spoken words perceived without word boundaries inserted, the graph on the left, the words with inserted WBs show somewhat more syllables with segmental errors, but the result is not statistically significant. For both final word boundary deletions and syllable insertions, however, the graphs in the middle and the right, the effect approaches significance with $p<.08$ using Fisher's exact probability. That is, words perceived with final word boundaries deleted, (cf. middle graph) have more syllables with perceptual errors than words perceived with no word boundary deleted. Similarly, on the right, words perceived with syllables inserted have more syllables with perceptual errors than words perceived with no inserted syllables. In other words, change in lexical structure of a perceived word does involve more perceived segmental errors in that word, as predicted.

In summary, the results of the analysis provide evidence for higher level processes introducing segmental errors in slips of the ear, where higher level processes can be simply identified by changes in lexical structure from the spoken to the perceived utterance.

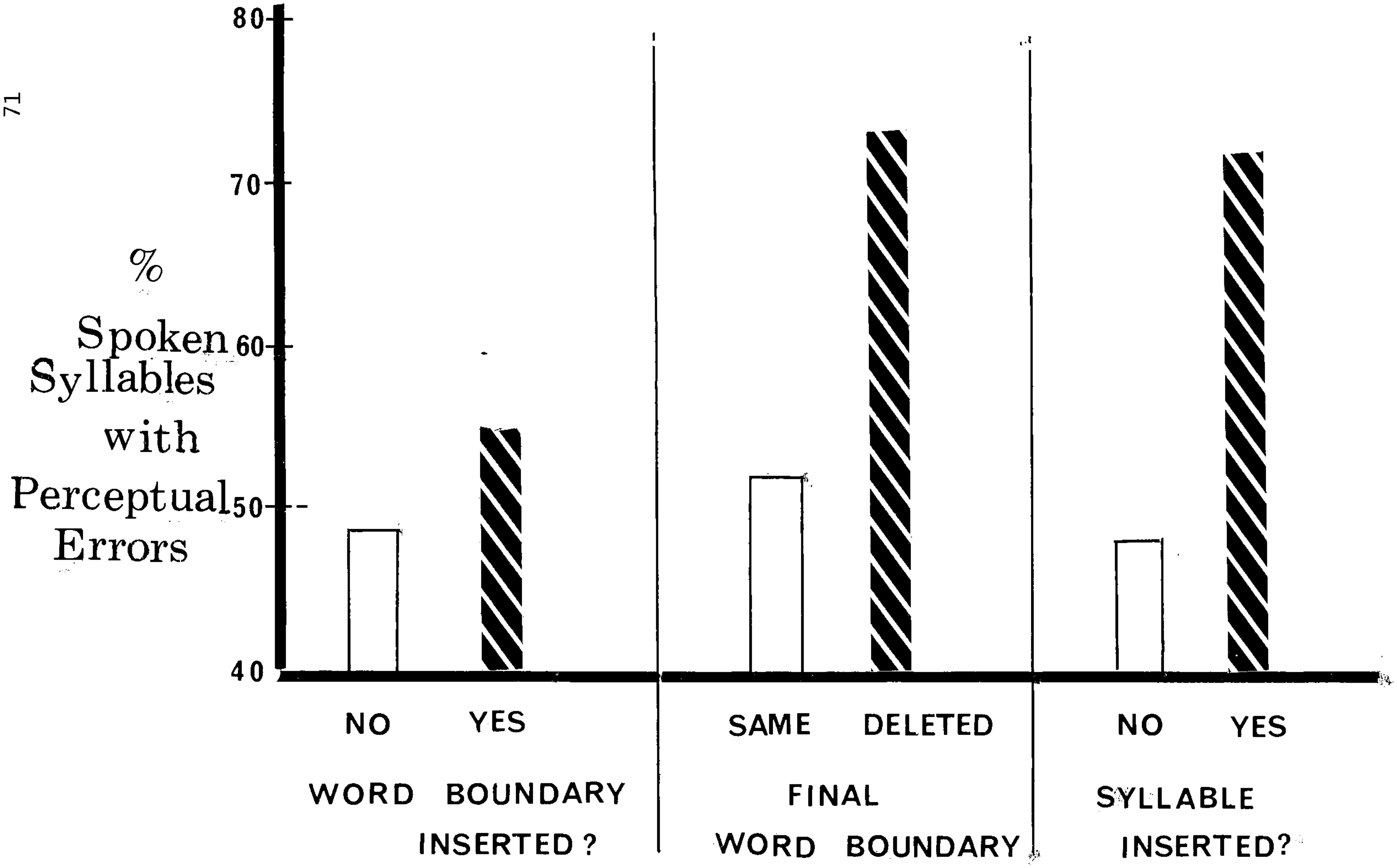


Figure 2. Types of Change in Lexical Structure in Perception of Spoken Words

What we Listen for

Louis Goldstein

[Paper presented at the 90th Meeting of the Acoustical
Society of America, San Francisco, California,
November, 1975]

Most currently accepted views of the speech understanding process posit at least two interactive components: first, a mechanism for sampling the acoustic signal, and secondly, a mechanism for making higher-level decisions about what word or phrase has been perceived. The second mechanism is generally assumed to have as input, not only a sampling of the acoustic signal, but listeners' expectations as well, (where expectations can be phonological, syntactic, semantic or extralinguistic). With respect to the interaction of these processes, it seems reasonable to ask whether there are certain aspects of the acoustic signal that are more important than others in this decision making. Put another way, parts of the acoustic signal might differ in the degree to which they constrain the decision mechanism to make certain choices. If different aspects of the acoustic signal *do* differ in this way,it seems a reasonable first hypothesis that those aspects that are most salient, most constraining, are those with the greatest energy. Thus, if one compares stressed with unstressed syllables, this view predicts that the stressed syllables should be more salient than the unstressed ones; and if one compares vowels with consonants, vowels should be more salient than consonants.

These predictions can be tested in a situation in which contextual information and expectancy are minimized. Such an experimental situation was designed for example, by Pickett and Pollack (1963). They spliced phrases out of spoken sentences and played them to listeners to identify. They found that in this situation, listeners make many errors when the spliced out segments are short (up to 800 msec).

By analyzing where listener's make most of their errors in a similar experimental situation, I attempted to test the hypothesis that the relative perceptual salience of certain aspects of the acoustic signal can be predicted by their relative acoustic energy. This hypothesis predicts, for example, that these listeners will make fewer errors in stressed syllables than in unstressed syllables and fewer errors involving the vowels than consonants.

Table 1: words and phrases used in experiment

DEEM
FAIN
YAWL
ROGUE
ZOOMS
TITHES
VOW
BIB
WED
CAM
GULL
NOOKS
MOTH
EEL
ALE
THONG
OAFS
LUTE
CHIVE
OWL
HICK
ELKS
PAP
SHUN
SOOT
JOT
YOKEL
SULFIDE
DILATE
SHODDY
CATCHY
VACUUM
GAZELLE
ROTUND
TYPHOON
MOTET
SUPINE
BEGUILE

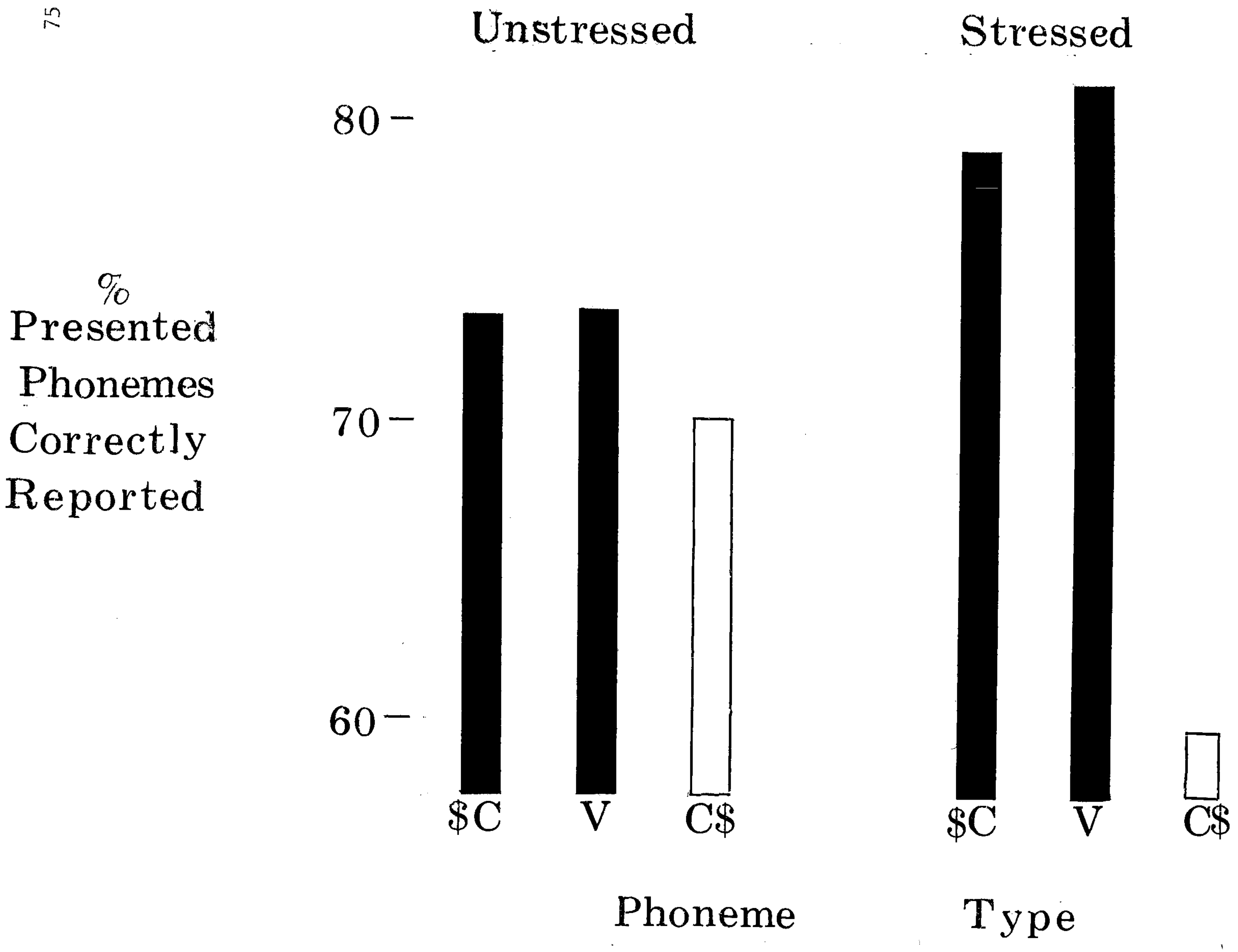
LOUSED YOU
FOOTFALL
BAIT THIEF
KNEE-TYPE
THIS CHOW
HOME BOUND
WILL ZIP
IT JIBES
CAUSED SOOT
CALL THUGS
BUILD MOUNDS
PUT SHIMS
NUMINOUS
RAVAGES
PULSATION
DEMENTED
REFEREE
GUARANTEE
FAUCET NUT
KETCHUP LUMP
HIS VISAGE
THOSE BOVINES
GETS SHAMPOO
YOU ADDUCE
WALKOUT DAY
WOODWIND SHOW
COULD LOG-JAM
THAT PAYOFF
ATE THE GAUZE
NOT HIS CHUMP
WE CHATTED
LOOKED JAZZY

The stimuli chosen for playing to listeners included 70 English words and phrases. They were chosen so as to be as phonetically balanced as possible, while at the same time controlling the word frequency of the stimuli. The frequency control was used so that differences in intelligibility could be due as much as possible to phonetic rather than expectation factors. Each stimulus included at least one word with a frequency of either one or two per million in the Kucera and Francis (1967) word count.

The 26 monosyllabic words were all stressed and each of the standard American English vowels (except ɔɪ) were included in two of these words. The initial consonant portion of these words were all single consonants - with each of the 22 English syllable initial consonants appearing in one word. Polysyllabic stimuli were chosen to include both single words and two and three word phrases. The single stressed syllable of each of the stimuli fell approximately an equal number of times in each syllable position. The initial consonant portions of all the syllables consisted of a single consonant, with all English consonants occurring at least once in stressed and unstressed syllables. In addition all American English vowels (except ɔɪ) occurred at least twice in stressed and twice in unstressed syllables. Syllable final consonants were left uncontrolled in the entire stimulus set, and include clusters as well as single consonants. A list of the stimuli is found in Table 1.

Sentences were then constructed in which these words and phrases were embedded, with the constraint that the last phoneme in the sentence *before* the stimulus and the first phoneme in the sentence *after* the stimulus were voiceless stops. This facilitated the editing process. The sentences were recorded by a speaker of general American English, naive to the purpose of the experiment. The stimuli were digitally edited out of the sentence and then rerecorded onto analog tape in a random order; the tape was then played, over a loudspeaker, to an introductory linguistics class. Students were told that the stimuli were words and phrases spliced out of content and were instructed to write down some English word or phrase for each of the stimuli; they were told to guess if they were unsure. Listeners written responses were then assigned a phonemic transcription, by me, using standard Southern California English as a reference dialect. Responses of 12 subjects were analyzed.

The listener's responses were analyzed so as to determine which of the phonemes of a given stimulus were correctly reported in the listener's response. Scoring proceeded as follows: First, the syllables of a stimulus were matched with the syllables of a given response in the following way. If the stimulus and response had the same number of syllables, then the first syllable of the stimulus was matched to the first syllable of the response,


Unstressed
Stressed
%
Presented
Phonemes
Correctly
Reported
80
70
60
$C
V
C$
$C
V
C$
Phoneme
Type

the second to the second, the third to the third. If the response involved the deletion or addition of a syllable(which was rare),the syllables were matched so as to maximize the number of phonemes common to the syllables being matched. *Vowels* were then considered to be correctly reported if the response syllable that was matched to the stimulus syllable included the stimulus vowel. An initial stimulus *consonant* was considered to be correctly reported if the response syllable that was matched to the stimulus syllable included this consonant in prevocalic position, or if the response syllable preceding the matched syllable included this consonant as its *final* phoneme. This criterion was the same for final stimulus consonants, except the opposite in direction.

Combined results are presented in Figure 1. There was very little difference in the percentage of correct reports for initial consonants compared to vowels. This is true in both stressed and unstressed syllables. Differences were found to be *not* significant using a Wilcoxon matched-pairs, signed ranks tests. However, a difference *did* emerge between stressed and unstressed syllables. More initial consonants of stressed syllables were reported correctly than of unstressed syllables ($p < .05$) and this is true for vowels as well ($p < .01$). While the final consonants show a rather different pattern,namely far fewer correct reports in stressed syllables than in unstressed. It must be remembered that the final consonants were not controlled for, and the final consonants of stressed syllables did include 3 times as many clusters as the final consonants in unstressed syllables, which may account for this reversal.

These results certainly do not support the hypothesis that relative perceptual salience can be predicted simply on the basis relative acoustic energy. However, stressed syllables do seem to be more salient than unstressed ones (at least for initial consonants and vowels). While this is an intuitively unsurprising result, one must ask why we get this effect. The explanation cannot lie in the fact that stressed syllables have more energy. than unstressed ones, since vowels have more energy than consonants, but are not more perceptually salient. Of course, the relationship between stressed and unstressed syllables is not *just* one of amplitude and duration, there are F0 differences, as well. One hypothesis that could account for the result, is that the complex of acoustic features that indicate stress (including energy, and F0) serve to direct the listener's attention to certain portions of the acoustic input. Such a hypothesis would be consistent with some experiments done (by Shayne and Gass) in the UCLA Phonetics Laboratory. These results indicated that subjects tend to report the location of clicks that are superimposed on a speech signal more accurately when they are on stressed syllables than when they are not, even when possibly confounding syntactic variables are controlled for. Thus, the relative perceptual salience of stressed syllables in the experiment may well be due to a hypothesized attention-focusing effect of stress.

References

Kucera, H. and Francis, W.N. (1967),*Computational Analysis of Present-day American English*. Providence, Rhode Island: Brown University Press.

Pickett, J.M. and Pollack, I. (1963), Intelligibility of excerpts from fluent speech: effects of rate of utterance and duration of excerpt. *Language and Speech, 6,* 151-171.

Shayne, J. and Gass, S. Effect of stress on click localization accuracy. unpublished manuscript, UCLA.

An Investigation of the Role of Stress as a Factor in Speech Perception

JoAnne Shayne and Susan M. Gass

[Department of Linguistics
University of California, Los Angeles]

INTRODUCTION

Several investigations have been conducted to discover the effect of certain factors on speech perception. Ladefoged and Broadbent (1960) ran an experiment in which they superimposed extraneous clicks on various syllables in sentences and on strings of numbers on the hypothesis that the segmentation of the stimulus utterance would prove to be larger than a single phonetic segment. The results of their experiment confirmed the hypothesis but left undefined the limits of this "larger unit". They found also that subjects tended to hear the click as occurring slightly before its objective position. This experiment was discussed further in Ladefoged (1967).

Fodor and Bever (1965), and Fodor, Bever, Garrett (1975) hypothesized that the undefined unit of perception corresponded to the grammatical constituent and that listeners would tend to hear the click as occurring at the boundary of the constituent within which it objectively occurred. Where constituents overlapped, the major syntactic boundary would exert the strongest "pull". They superimposed extraneous clicks on sentences which had overlapping syntactic boundaries and found that their hypothesis was also confirmed. The major syntactic break did appear to be attracting the click.

Another factor which may prove to have significant influence on speech perception, and which has not yet been similarly investigated, is the effect of stress. The results of the two experiments below will show that listener perception tends to be drawn toward the positon of stress.

EXPERIMENT I

The stimuli in this experiment consisted of high frequency beeps, approximately 67 milliseconds in length, superimposed on twelve sentences. All the sentences consisted of the words:

That he was happy was evident from the way that he smiled.[1]

This sentence was spoken in three different ways. In four of the twelve sentences there was an emphatic stress on the first occurrence of the word "was"; four others had an emphatic stress on the first syllable of "happy"; and the remaining four had an emphatic stress on the first syllable of "evident." In each group of four sentences the superimposed beep occurred in a different place on each occasion. The four different locations for the beep were the three syllables "was" "hap(py)" and "ev(ident)" described above, and the word "from." The twelve sentences were randomly ordered on a tape.

The tape was played for thiry-six subjects, all of whom had printed copies of the sentences in front of them. The subjects were asked to listen to each sentence and then to mark, by means of a vertical slash above the proper place, where they thought the beep occurred. This could be either on or between syllables. They were told to mark every sentence once (and only once), even if they were not sure they had heard a beep at all.

EXPERIMENT II

This experiment consisted of ten "sentences", each composed of nine nonsense syllables (CV structure). A high-frequency beep, approximately 67 milliseconds in length, was superimposed on one syllable in each sentence. The sentences were divided into five pairs; within each pair, the beep occurred on the same syllable, while the position of stress varied. The sentences were then randomly ordered. The basic sentences read as follows:

1) dera bolimo nu gahipe
2) tikape rosu nala botu
3) sena padi kerulema ni

(word boundaries strictly arbitrary)

The tape was played for thirty-six subjects, all of whom had printed copies of the sentences in front of them. The instructions were the same as those in experiment I.

RESULTS

Experiment I.

Figure 1 illustrates the responses to the twelve sentences of experiment 1. The sentences have been grouped according to the distance between the beep and the stress. Table 1 summarizes these data, bearing in mind that subjects' responses can occur on or between any syllables.

Table 1. Mean number of responses to beeps at different distances from the stressed syllable.

Number of possible response positions between beep and stressed syllable	Correct responses	Incorrectly placed on stressed syllable	Incorrectly placed on other syllables
0 (same syllable	16	--	19.3
2 (next syllable)	11.5	11	14
4 or 5 apart	2	14	18.7
7 apart	3	9.5	23.5
9 or more apart	1.5	10	24.5

Experiment II.

Figure 2 illustrates the distribution of responses to the ten sentences of experiment II. The mean number of correct responses to the sentences in Group A (those in which beep and stress are on the same syllable) is 28; the mean number of correct responses to the sentences in Group B (those in which stress and beep are on different syllables) is 11. A chi-square test was run which showed the difference between the number of correct responses to both groups of sentences to be significant at the .025 level (4 degrees of freedom). Chi-square = 11.18.

DISCUSSION

The combined results of the two experiments present strong evidence that the location of stress does influence where subjects tend to hear the beeps. The sentences in which both stress and beep occur on the same syllable present the most clear-cut examples of this (see Figures 1 and 2, A sentences).

In order to investigate the results in some detail, the two experiments will be discussed separately. Figure 1 (experiment I) shows that the number of correct responses decreases as the distance between beep and stress is increased, with the subjects consistently tending to favor the stressed syllable as the correct response.

Comparing sentences 1, 2 (Figure 1, Group B) and 3 (Group C) to sentences 1 and 2 (Group C), it appears that when there is a distance of no more than 5 positions between beep and stress,^{2} and both occur on the same side of the major syntactic boundary, there tends to be a clustering of responses around stress, but it is not overwhelmingly on the stressed syllable itself. However, given the same distance limitation with beep and stress on *opposite* sides of the major syntactic boundary, there is a marked preference for the stressed syllable itself as the correct response. This indicates that the beep has more pull on listener perception when it occurs on the same side of the major syntactic boundary as the stressed syllable, although the stressed syllable still draws a significant number of responses.

Looking at the sentences in groups D and E (Figure 1) where beep and stress are separated by 7 or more positions and occur on opposite sides of the syntactic boundary, it is clear that there is still a definite clustering around the stress, but no significant preference for the stressed syllable itself. This suggests further experiments to discover at what point stress begins to lose its pull on listener perception and at what point it no longer has any influence.

The results of experiment II confirm those of experiment I in that when beep and stress occur on the same syllable (Figure 2, Group A) there is an overwhelming number of correct responses. The responses for the remaining sentences (Figure 2, Group B), however, produce results which are somewhat inconsistent with those of similar sentences in experiment I[3]. Instead of clustering around the stressed syllable, there is a clustering of responses around the beep, although there is a large number of incorrect responses occurring on the stressed syllable. Since the number of positions between stress and beep in these sentences places them within the same group as those in Figure 1, Group C, distance cannot be considered to be a significant factor in the discrepancy. One possible explanation is that whereas the sentences of experiment I were spoken as they would be in conversation, those of experiment II were more slowly and carefully articulated, producing a clearer distinction between one syllable and the next. Such a reading may have given the listener a better opportunity to hear on which syllable the beep actually occurred. This, too, is an area which merits further investigation. The ease with which the listener is able to distinguish between one syllable and the next may prove to be an important factor in perceptual processing.

However, it is more likely that the discrepancy can be attributed to the fact that the subjects were unable to segment the nonsense syllables of experiment II into any kind of grammatical or lexical categories and were thus forced to process the sounds in terms of the individual syllables. If so, the results of these experiments, in addition to confirming our original hypothesis, have also provided evidence in support of the findings of Fodor and Bever.

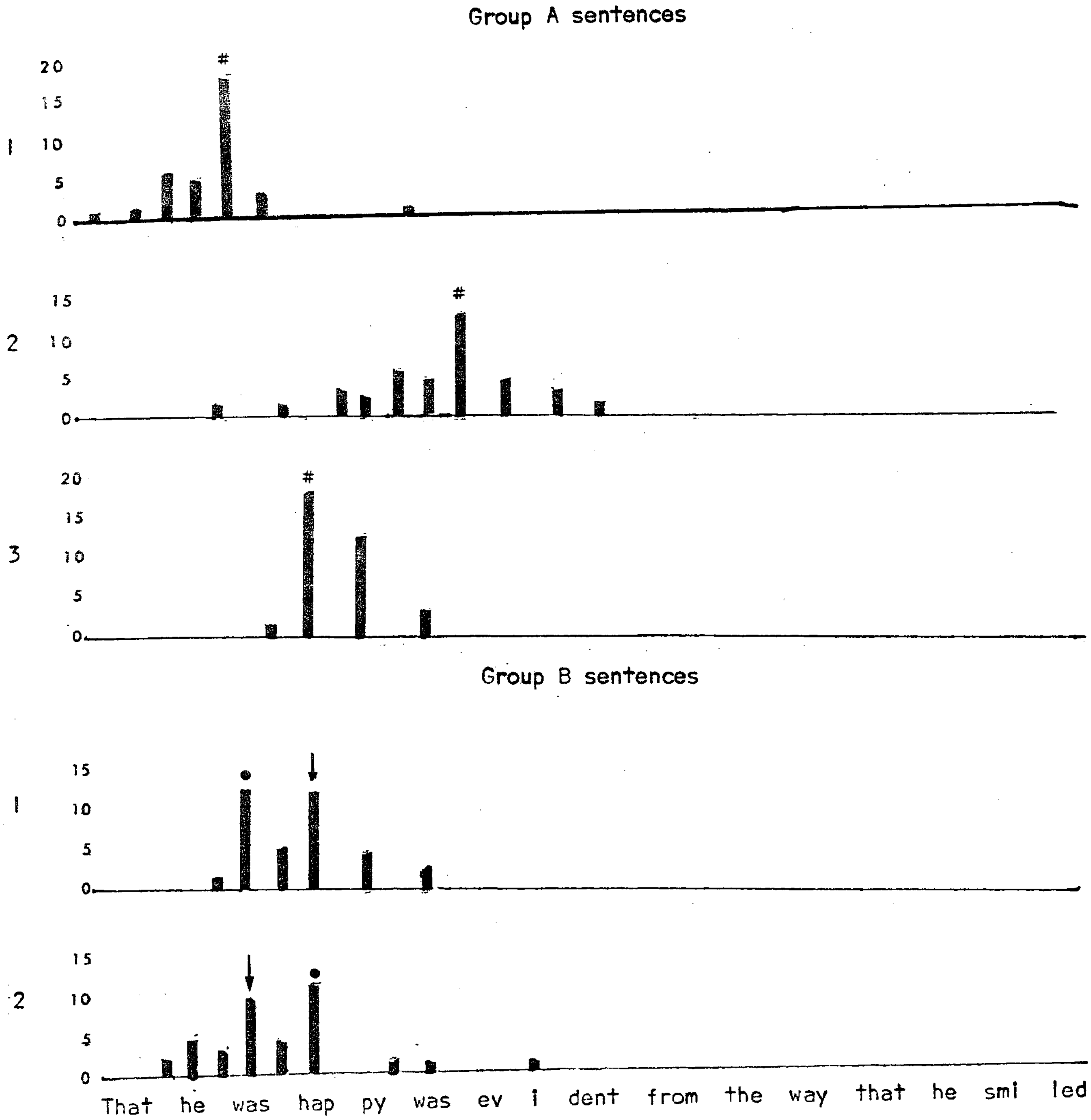
Footnotes

1. This is the sentence used by Fodor and Bever (1965).

2. Because we have no examples of sentences in which beep and stress are separated by six positions, we are forced here to assume that five positions is the cut-off point.

3. Since there are no syntactic boundaries in the sentences of experiment II, we will assume that the "similar" sentences of experiment I are those in which beep and stress occur on the same side of the boundary.

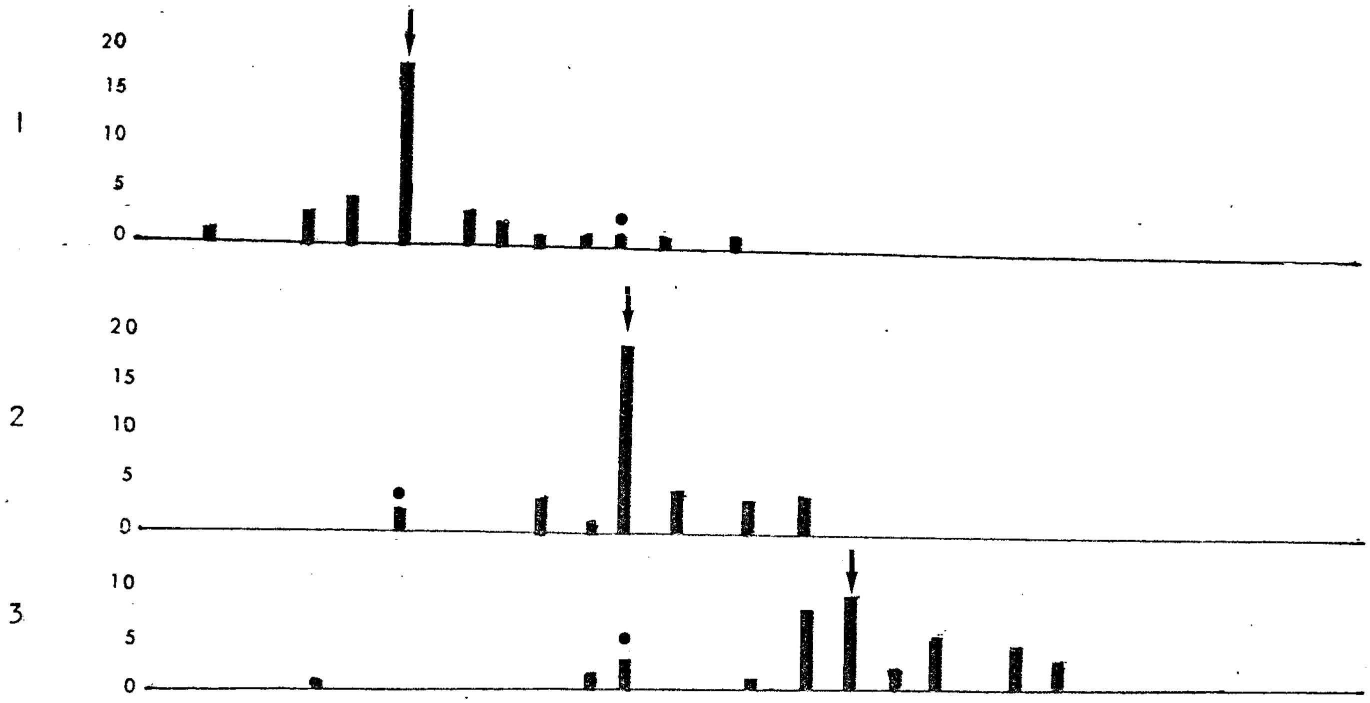
Figure 1. Results from experiment 1.



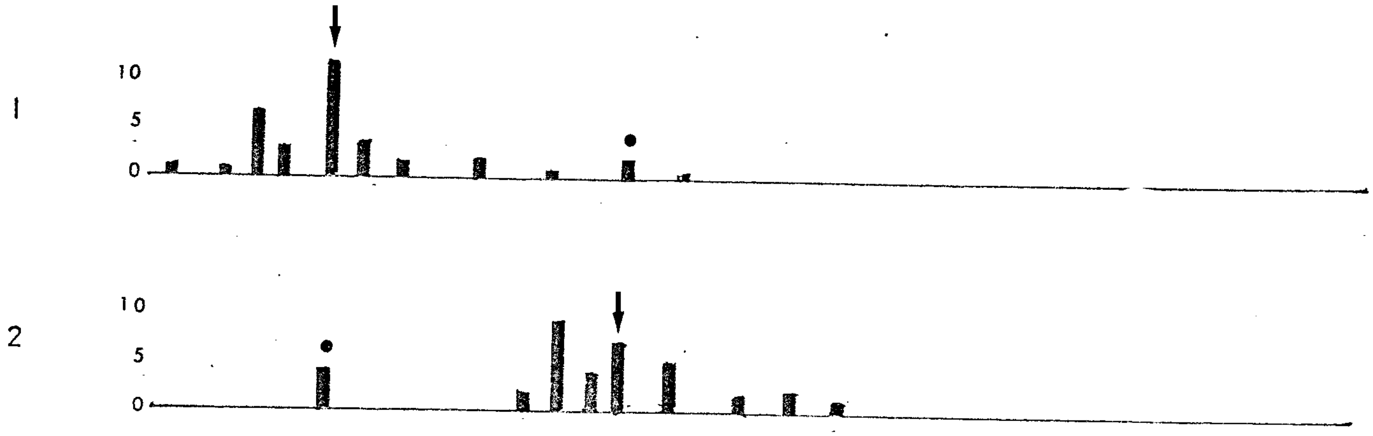
Group A- beep and stress on same syllable
Group B- beep and stress two positions apart
Group C- beep and stress four or five positions apart
Group D- beep and stress seven positions apart
Group E- beep and stress nine or more positions apart

\# - beep and stress on same syllable
→ - stress
● - beep

Group C sentences



Group D sentences



Group E sentences

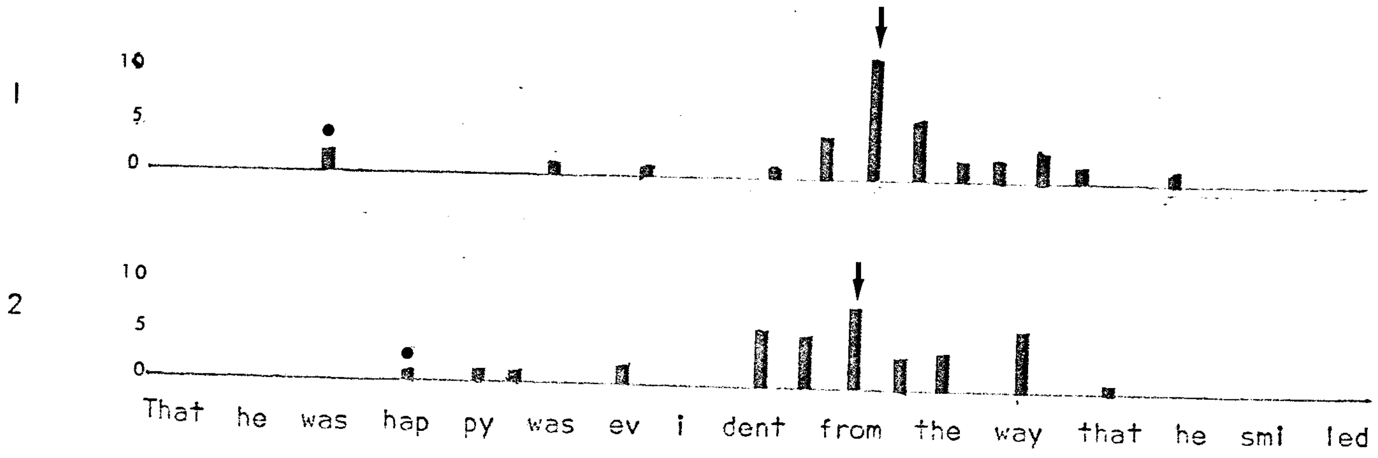
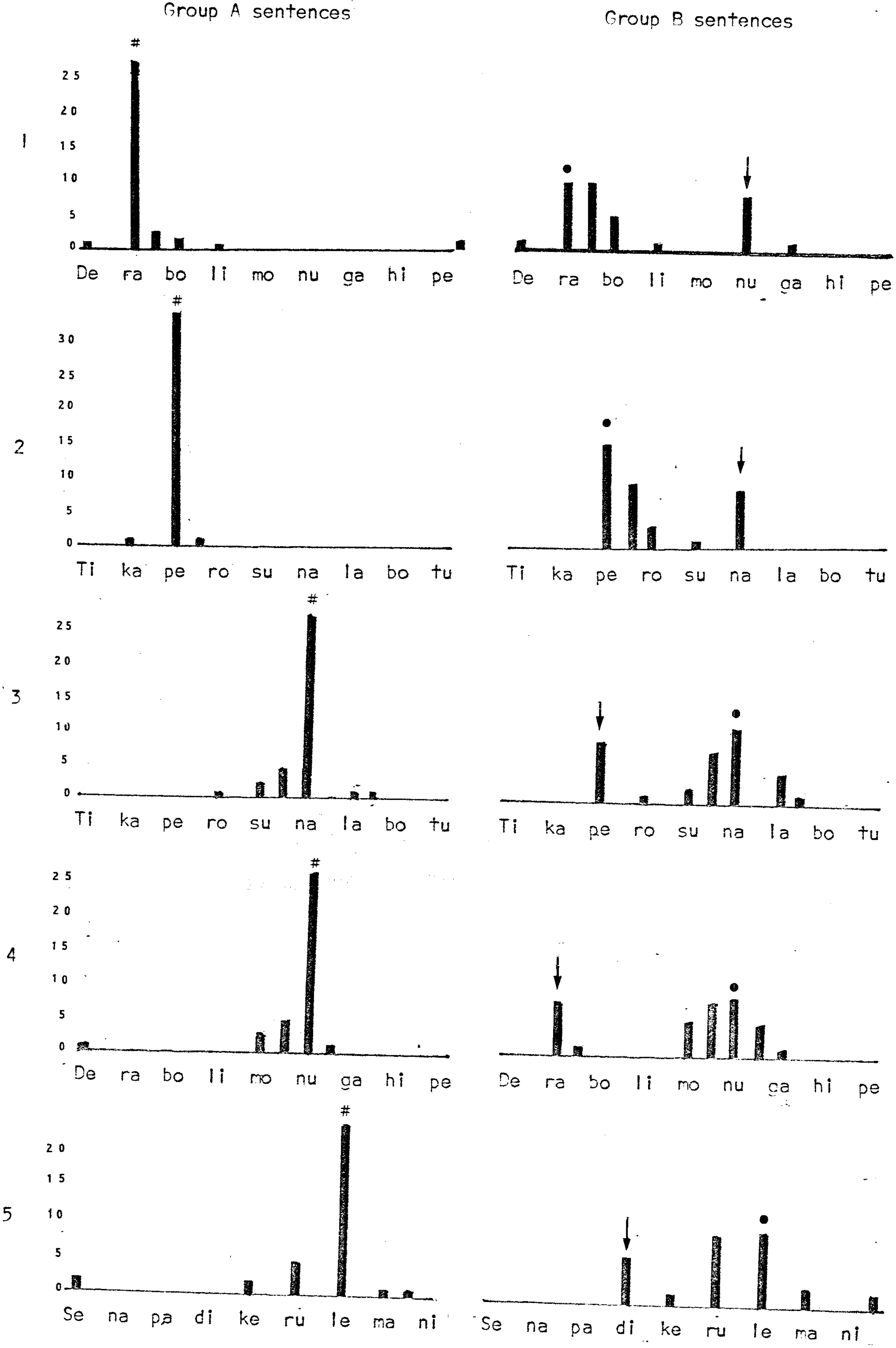


Figure 2. Results from experiment II



Acknowledgement

We would like to thank Ian Maddieson for helpful suggestions. We especially thank Peter Ladefoged for direction, correction and infinite patience.

Bibliography

Fodor, J. A. and T. G. Bever. "The Psychological Reality of Linguistic Segments" *Journal of Verbal Learning and Verbal Behavior* 4:414-420. 1965.

Fodor, J. A., T. G. Bever and M. F. Garrett. *The Psychology of Language* McGraw-Hill Book Company, New York. 1974.

Ladefoged, Peter. "Units in the Perception and Production of Speech" *Three Areas of Experimental Phonetics* Oxford University Press, London pp. 143-172. 1967.

Ladefoged, P. and D. E. Broadbent. "Perception of Sequence in Auditory Events" *Quarterly Journal of Experimental Psychology* 12:162-170. 1960.

Sex, Language, and the Brain, Part I: A Review of the Literature on Adult Sex Differences in Lateralization [1,2]

Richard A. Harshman and Roger Remington[3]

[Phonetics Laboratory, Department of Linguistics,
University of California, Los Angeles]

Abstract

While it is known that men and women differ in relative performance on certain verbal and spatial tasks, the cause of this difference has not yet been determined. There is evidence, however, that both cultural and biological factors are involved. One possible biological factor is brain organization.

A review of clinical and experimental data indicates that the brains of men and women tend to show differences in hemispheric specialization for some (but probably not all) language and spatial functions, with men showing a greater average degree of lateralization than women. This conclusion is supported by several different types of evidence: clinical studies of the effects of unilateral brain damage or commissurotomy on verbal and spatial tasks; anatomical studies of cerebral asymmetries in language related cortical areas; electrophysiological studies comparing right and left hemispheres during stimulation or behavioral tasks; tachistopic perceptual studies of right and left visual field advantages to verbal and visuo-spatial stimuli; and other, indirect evidence. Data on lateralization of children must be considered separately, because of complexities in interpretation arising from boys' developmental lag relative to girls.

Introduction

Are men's and women's brains organized differently for the processing of language? Several lines of evidence point to this possibility. Sex differences in verbal and nonverbal *performance* have been noted for many years: women consistently outperform men on tests of verbal fluency, while men do better than women on certain visuo-spatial nonverbal tasks (for reviews, see Bock, 1973; Garai and Scheinfeld, 1968; Maccoby and Jacklin, 1974; Oetzel, 1966; Tyler, 1965, Ch. 10). There is evidence that both cultural *and* biological factors contribute to these performance differences (e.g., see discussions in Bock, 1973; Buffery and Gray, 1972; Dawson, 1972; Maccoby and Jacklin, 1974). However, it is the evidence for biological factors which is of particular interest here. One source of such evidence is the data linking sex hormone levels and androgen-estrogen balance to individual differences in verbal and/or visuo-spatial performance (Bock, 1973; Broverman and Klaiber, 1969; Dawson, 1972; Masica, Money, Ehrhardt, and Lewis, 1969; Petersen, 1973). Another, more indirect source of evidence is the data indicating genetic control of one or more of the factors underlying cognitive sex differences. In particular, there is evidence from family studies that a sex-linked recessive gene may account for much of the higher mean spatial performance found in males (Bock and Kolakowski, 1973; Hartlage, 1970; Stafford, 1961; Yen, 1973). Further suggestive evidence for genetic influences comes from examination of the different patterns of verbal vs. spatial performance found in individuals with too few or too many sex chromosomes, i.e., XYY, XO (Turner's syndrome), and XXY (Klinefelter's syndrome) (Buckley, 1971; Garron and Vander Stoep, 1969; McKerracher, 1971; Money, 1964, 1968).

Whatever the nature of the biological factors contributing to verbal and spatial sex differences, it seems likely that these factors act by directly or indirectly influencing the organization or functioning of the brain. Yet almost nothing is known about the different patterns of brain functioning which might contribute to the different patterns of abilities found in normal persons. Nonetheless, there is evidence from many lines of research which indicates that there are, in fact, sex differences in the neural organization of verbal and visuo-spatial functions in the human brain. This article presents a survey of clinical, anatomical, electrophysiological, tachistoscopic perceptual, and other data on the relative degree of left cerebral hemisphere specialization for language (and right cerebral hemisphere specialization for spatial abilities) in men vs. women. It will be proposed that the pattern of findings across these various types of data indicates a greater mean degree of hemispheric specialization in adult men than in adult women. It is further suggested that such a neuroanatomical difference would be consistent with the observed differences in verbal and visuo-spatial performance of men and women, and would provide a biological basis for one or more of the factors which underly these behavioral sex differences. Not only could such a neurological-performance link add to our knowledge about average differences between men and women, but it might also suggest useful hypotheses about the relationship between brain mechanisms and linguistic and visuo-spatial abilities.

A Pattern of Evidence for Sex Differences in Brain Lateralization

While it has been suggested by a number of authors (e.g. Bakan, 1971; Buffery and Gray, 1972; Ehrlichman, 1971; Gur and Gur, 1974; Kimura, 1969; Lansdell, 1962; Levy, 1973; McGlone and Kertesz, 1973; Sherman, 1974; Wada, 1974) that men and women have different patterns of functional specialization of the two cerebral hemispheres. This suggestion has usually been quite tentative, based on only a few pieces of corroborating data. Further, there has been some disagreement as to the nature of this sex difference. With one exception, the authors cited above have suggested that males show a greater "lateralization" (i.e., localization of a psychological function into a single hemisphere) for verbal and/or visuo-spatial skills than females. According to such a theory, males would have the left hemisphere more specialized for verbal, and the right hemisphere more specialized for visuo-spatial skills, whereas females would tend to have both skills more equally represented in both hemispheres.

Developmental differences

Buffery and Gray (1972), however, have proposed the opposite theory, i.e., that females are more lateralized for both language and spatial skills. But Buffery and Gray based their theory mainly on a review of developmental data from children. Their primary conclusion is that language lateralization proceeds more rapidly for girls than for boys, and thus girls will show a more advanced level of lateralization at any given age. This *developmental* theory is fairly well supported (see for example, Witelson and Pallie, 1973; and evidence reviewed in Buffery and Gray, 1972), and is consistent with the fact that girls seem to show a generally more rapid behavioral and physical maturation than boys (Ounsted and Taylor, 1972; Scheinfeld, 1944; Sinclair, 1969; Terman and Tyler, 1954), including a more rapid cerebral maturation (Taylor, 1969). But the additional conclusion of a basic tendency to greater female adult lateralization need not follow from this developmental data. As Buffery and Gray themselves point out, "sex differences in children are difficult to interpret when there is an advantage in favor of girls, since this may always be due to their general maturational advance over boys" (Buffery and Gray, 1972, p. 131).

The same problem is encountered if one compares the physical stature of boys and girls matched on the basis of chronological age. Despite the eventual tendency for males to be larger, the two sexes are very similar in height and weight throughout most of childhood, and, in later childhood, girls tend to be even taller and heavier than boys (Breckenridge and Vincent, 1965; Mussen, Conger, and Kagan, 1963; Sinclair, 1969; Tanner, 1970). At any given chronological age, the boys' less advanced physical development compared to girls masks an underlying tendency to larger physical size. It seems plausible that a similar masking phenomenon might complicate the comparison of the degree of lateralization between boys and girls. In this regard, it is interesting to note some suggestive findings by Marcel, Katz, and Smith (1974). They found that when boys and girls were matched on level of development of reading ability (rather than chronological age), boys showed greater tachistoscopic visual field asymmetries (right field superiority for verbal stimuli) than did girls. Although a number of other factors may be involved, one possible interpretation is that it is *boys* who would show the greater verbal lateralization if only an appropriate non-chronological basis could be found for matching them with girls.

While some aspects of lateralization might be present from birth (e.g., see Molfese, 1973; Witelson and Pallie, 1973), others appear to show clear developmental trends (for discussion, see Brown and Jaffe, 1975; Buffery and Gray, 1972; Harshman and Krashen, 1974; Krashen, 1972; Krashen and Harshman, 1972a, 1972b). Indeed, different aspects of language processing might show different developmental curves (Harshman and Krashen, 1974) and different patterns of sex differences (see Harshman, Remington, and Krashen, 1975).

Figure 1 shows hypothetical developmental curves for lateralization of some function not fully lateralized at birth. These curves are consistent both with the evidence for greater female lateralization during childhood (e.g. cited by Buffery and Gray, 1972) and the evidence for greater male lateralization in adulthood (cited in this article, see below). Three very simple assumptions were made in drawing these curves: (a) that boys and girls have the same basic pattern of development, thus the curves have the same overall shape; (b) that boys have a slower rate of development, thus their curve is expanded slightly to the right (extended in time); and (c) that boys have a larger underlying degree of lateralization at an equivalent developmental stage, thus their curve is drawn to a slightly greater vertical scale. Of course, the real patterns of development of lateralization for boys and girls are probably more complicated, and may differ for the two sexes and for different types of linguistic (and non-linguistic) functions. (It is even possible that female hemispheric asymmetry *decreases* in late childhood due to continuing development of right hemisphere language capacities).

Adult Differences

Because of the difficulties in interpreting childhood sex differences in lateralization, it is necessary to concentrate on *adult* lateralization data when considering the question of sex differences in brain organization for language. Lateralization studies which compare men and women do not always reveal a statistically significant sex difference (perhaps because of the particular variables measured, sampling fluctuations, or "noise" in the data). However, in the studies we have seen(reviewed below), a clear pattern seems to emerge: when sex differences in lateralization are found for adult subjects, they are almost invariably in the direction of greater *male* lateralization (Table 1).

The evidence for greater male lateralization in adulthood comes from four basic sources: clinical studies of the effects of unilateral brain damage or commissurotomy on verbal and spatial tasks; anatomical studies of cerebral asymmetries in language-related cortical areas; electrophysiological studies comparing right and left hemispheres during sensory stimulation or behavioral tasks; and tachistoscopic perceptual studies of right and left visual field advantages to verbal and visuo-spatial stimuli. There are also other data of several kinds providing indirect support for greater male lateralization.

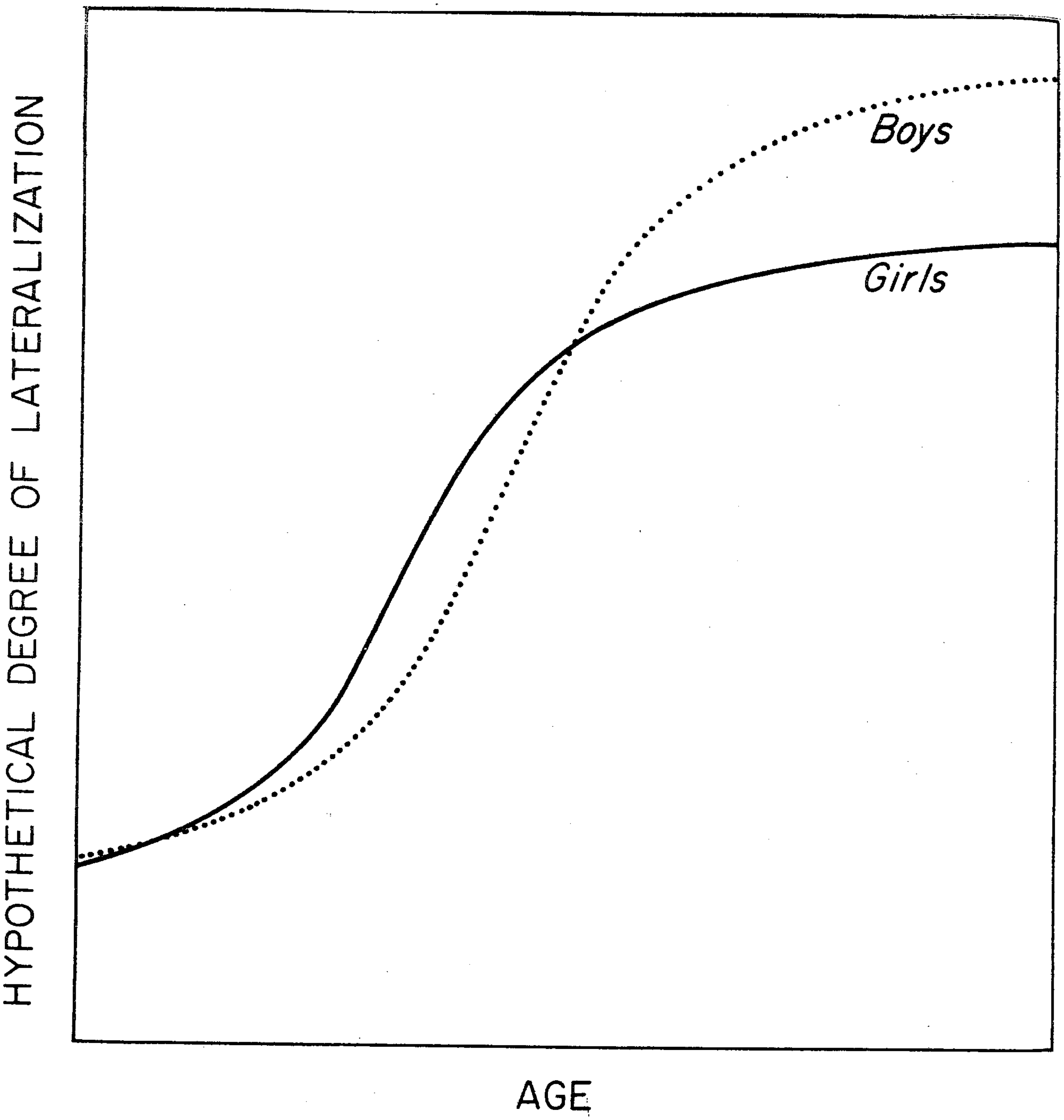


Figure 1. Simplified hypothetical curves for the development of some aspects of lateralization in boys and girls: both curves have a common shape (arc-tangent) but the girls' time scale is contracted (.9x) to represent faster development, and the boys' vertical scale is expanded (1.25x) to represent a basic tendency to greater lateralization. The females' faster rate of development causes younger girls to appear more lateralized than boys.

Table 1

Adult Sex Differences in Lateralization: Distribution of Main Findings According to Type of Evidence Involved[a]

	Sex Showing Greater Lateralization		
Type of Evidence Involved	Males	Neither	Females
Verbal Lateralization			
Clinical Data	2	2	0
Anatomical Data	1	0	0
Tachistoscopic Data	2	2	0
Visuo-spatial Lateralization			
Clinical Data	3	0	0 (or 1?)[b]
Tachistoscopic Data	3	2	0
General Lateralization[c]			
Clinical Data	1	0	0
Electrophysiological Data	3	0	0
Anatomical Data	1	0	0

[a] Studies are those described in text; indirect evidence and unpublished findings not counted

[b] Depending on the interpretation of Lansdell (1968a); see footnote 4

[c] Evidence of lateralization which is not specific to verbal or visuo-spatial functions.

Clinical data.

Using commissurotomy patients, Bogen, DeZure, TenHouten, and Marsh (1972) found that the ratio of verbal to spatial test performance of the left hemisphere was significantly greater for the males than for the females tested, suggesting greater linguistic specialization of the left hemisphere in males. Similar results have been found in tests of brain-lesioned patients: adult males show a greater and more selective impairment of verbal skills after unilateral damage to the left hemisphere (Lansdell, 1961, 1973) amd a similarly greater and more selective impairment of nonverbal visuo-spatial skills after unilateral damage to the right hemisphere (Lansdell, 1962, 1968b; McGlone and Kertesz, 1973). Although complicated by other factors, this result is also suggested by data in Lansdell and Urbach (1965). (Bakan (1971) also reports that analysis of clinical data indicates greater male lateralization, but he provides no details; Sherman (1974) provides several citations of published and unpublished data to support the same conclusion, but does not give details of her statistical analyses).

At other times, however, these sex differences are not obtained with clinical populations or are merely reflected in trends which do not approach statistical significance: Lansdell (1968a) and McGlone and Kertesz (1973) failed to find statistically significant sex differences in verbal lateralization; Lansdell and Urbach (1965)failed to confirm the pattern of spatial sex differences in lateralization reported by Lansdell (1962). Finally, other data reported by Lansdell (1968a) show a pattern of sex differences which is difficult to interpret.[4]

The overall pattern in the clinical data indicates greater male lateralization of both verbal and visuo-spatial functions. However, sex differences seem to be often hard to detect or replicate using clinical methods, perhaps due to subtleties of the differences or difficulties inherent in clinical-experimental methodology. It has also been suggested that the sexes may differ in lateralization of some functions, but not others (Harshman, Remington, and Krashen, 1975; Lansdell, 1973). In light of such suggestions, it is worth noting that tests of proverb interpretation (Lansdell, 1961) and word association identification (Lansdell, 1973) have shown clinical evidence of sex differences in lateralization,while a vocabulary test (Lansdell, 1968a) and an aphasia battery (McGlone and Kertesz, 1973) have not. (However, the aphasia battery showed a nonsignificant trend in the direction of greater male lateralization).

Anatomical data.

Wada (1974) reported sex differences in the hemisphere asymmetry of physical brain measurements. When the size of left hemisphere regions in the vicinity of the classical "speech areas" (Left temporal planum and frontal operculum) was compared with the size of the corresponding areas on the right side of 100 adult human brains, there was a strong tendency for the left temporal planum (but not the frontal operculum) to be larger. This same tendency has been reported by several other researchers (see

Geschwind, 1974; Witelson and Pallie, 1973), but Wada noted, in addition, that female brains showed a significantly less consistent pattern of asymmetry than did male brains. Lansdell and Davie (1972) also noted neuroanatomical sex differences suggestive of greater male lateralization.

Electrophysiological data.

Experiments by Eason and others (Eason, Groves, White, and Oden, 1967; Eason, Groves, and Bonelli, 1967) have compared the average evoked response (AER) of the right and left occipital lobe to repetitive light flashes under two conditions: when the flashes are restricted to the right vs. the left visual field. These conditions show a different pattern of responses in right-handed, but not in left-handed, male subjects. For right-handed males, the relative amplitude of the responses in the two hemispheres varies with the side of visual field stimulated. For left-handed males, on the other hand, the amplitude is always largest in the right occipital lobe. The pattern of responses in females resembles those of left-handed males, i.e., regardless of handedness, females always show a larger evoked response in the right hemisphere (Culver, Tanley, and Eason, 1970; Eason, Groves, and Bonelli, 1967).

Another type of EEG evidence involves procedures which have recently been developed for measurement of an individual's degree of verbal or spatial lateralization. These procedures involve comparison of right and left hemisphere EEG activity during verbal vs. musical tasks (Herron, 1974) or during speech (Johnson, 1973). Comparison of lateralization scores for a group of males (Herron, 1974) with scores for a group of females (Johnson, 1973) reveals a higher incidence of left hemisphere language lateralization in the male group (Johnson and Herron, personal communication, 1974). Since different electrophysiological techniques are involved for the two groups, these results are only suggestive. However, the difference between the groups was also confirmed with a dichotic listening measure of asymmetry, and the dichotic scores correlated highly with the electrophysiological scores in both cases.

In a study still in progress at the UCLA Brain Research Institute, Brown, Marsh, and Smith (personal communication, 1975) are finding evidence of greater male lateralization of differences in processing of nouns vs. verbs. They are comparing average evoked responses within each hemisphere across two stimulus conditions: the word "fire" in "ready, aim, fire" and the identical physical stimulus embedded in the sentence "sit by the fire." The two conditions produce greater differences in shape of the AER in the left hemisphere than in the right, but this functional asymmetry is much less pronounced in women subjects.

Tachistoscopic perceptual data.

Males have been found to show greater visual field asymmetries than females on tachistoscopic spatial tasks (Kimura, 1969, 1973; McGlone and Davidson, 1973), indicating greater spatial lateralization. At other times, however, no significant sex differences have been observed on what seem to be similar tasks (Ehrlichman, 1971; Kimura, 1969). Two studies have tested language-related sex differences in lateralization using tachistoscopic *verbal* stimuli. In two of these studies (Ehrlichman, 1971, Marshall and Holmes, 1974), significant sex differences were found indicating greater male verbal lateralization. In the other two (Bryden, 1965; Marshall, Newcombe, and Holmes, 1973), no significant sex differences were observed.

Other data.

There is indirect evidence from several sources which suggests that males are more lateralized than females. (a) Women are reported to more frequently have trouble telling left from right (Bakan and Putnam, 1974; Wolf, 1973), a difficulty which has been linked to bilateral symmetry of the nervous system (Corballis and Beale, 1970). (b) In a review of five audiometric surveys covering sizeable populations, Kannan and Lipscomb (1974) find that males, but not females, show what might be called an "ear dominance," i.e., very slight superiority (less than 1 db) of the right ear for the detection of threshold level tones. (c) Ehrlichman (1971) suggests that the greater statistical independence (lower intercorrelation) for certain groups of ability test scores in males (Dye and Very, 1968; McFarlane, 1925; Gardner, Holzman, Klein, Linton, and Spence, 1959; Very and Iacono, 1970) might indicate greater physiological independence of their neural substrates, e.g., greater lateralization. (d) Gur and Gur (1974) measured the correlation between susceptibility to hypnosis (which, they hypothesized, requires "right hemispheric" nonanalytical, imaginative thinking) and the frequency of right lateral eye movements when pondering reflective questions (which has been linked to the use of "left hemispheric" analytical, verbal thinking; see Bakan, 1969; Kinsbourne, 1972, 1974; Weiten and Etaugh, 1973, 1974). As predicted, a significant negative correlation was found, which was strongest in right handed males. This correlation was reduced significantly in females and left handed subjects, indicating a greater frequency of reduced or reversed lateralization in these two groups[5].

It has also been suggested, however, that there is indirect evidence in support of the opposite conclusion, i.e., Buffery and Gray (1972) have suggested that the higher incidence of left handedness and dyslexia found among males would only be expected if males were *less* lateralized than females. This argument makes the plausible assumption that if one sex is naturally less lateralized, then it should, as a consequence, be more prone to those conditions which have sometimes been found to correlate with reduced lateralization. It seems likely, however, that the pathologically reduced lateralization sometimes found in dyslexics is the result of a malfunction of a developmental trend toward fuller lateralization on the part of the organism, whereas the nonpathological reduced lateralization found more frequently among females is the result of a different but healthy form of brain organization, a form of organization in which reduced lateralization is functionally appropriate, and, perhaps, one in which reduced lateralization would even have beneficial consequences (see Part III (Harshman and Berenbaum, 1975) for further discussion).

Questions also arise concerning the interpretation of sex differences in the incidence of left handedness. Ehrlichman (1971) points out that the slightly higher rate of left handedness found among males may be a cultural artifact, and that mixed hand-eye dominance is actually more frequent among females (for other data on greater female incidence of mixed hand-eye dominance see Kimura, 1969; Porac and Cohen, 1969). Furthermore, the relationship between handedness and cerebral dominance is not a simple one-to-one corresondence, so that the more frequent occurrence of reduced lateralization among left handers need not necessarily imply a more frequent occurrence of left handedness among those females who show reduced lateralization. Clearly, further study is needed to unravel the relationships among these factors. In this regard, it is noteworthy that Gur and Gur (1974) found both left handedness and femaleness to be independently associated with reduced lateralization, and that, in their data, the effects of these two variables were apparently additive--left handed females showed the most evidence of altered lateralization. (The implications of the handedness and dyslexia data will be discussed in more detail in Part III (Harshman and Berenbaum, 1975)).

There is one important source of evidence on sex differences in lateralization which still has not been considered: dichotic listening data. Discussion of dichotic listening has been deferred to Part II (Harshman, Remington, and Krashen, 1975) because a simple literature review of dichotic listening studies would not be adequately informative. There are only a few published studies which provide data on the relative lateralization of adult men and women, and their results are inconclusive; further data is needed before the implications of dichotic listening evidence can be properly evaluated. Such additional data is provided by Part II which reviews the published dichotic evidence on adult sex differences and also presents new analyses, by sex, of three additional dichotic data sets. These combined results will indicate that males show greater lateralization than females for perception of dichotic consonant-vowel monosyllables, thus further confirming the general pattern of greater male lateralization for some linguistic functions. However, comparison of these results with those reported for multiple pairs of dichotic digits and other stimuli will tentatively suggest that sex differences in lateralization may not be found when the dichotic task involves a heavy load on verbal short term memory. This again raises the possibility that only some lateralized functions will show sex differences in degree of lateralization,while others may not.

Summary and Conclusion

The available evidence for and against sex differences in lateralization has been reviewed. This evidence includes a comparison of normal men and women in terms of asymmetries in perception, brain electrical activity, and post mortem measurements of gross brain anatomy. In addition, men and women with unilateral brain damage are compared in terms of performance on standardized verbal and visuo-spatial tasks. The evidence reviewed has been gathered from over 30 published studies, plus some unpublished findings and work currently in progress at several laboratories. While no one study, considered by itself, provides conclusive evidence, the many studies, when considered in relation to one another, exhibit a convergence of findings which is hard to dismiss. Despite the numerous different sources of data and varied experimental techniques, there emerges a consistent pattern of evidence for greater adult male lateralization of at least some verbal and visuo-spatial functions. Two qualifications should be noted, however: (a) girls may sometimes exhibit greater average lateralization than boys of the same age, presumably because of their faster rate of physical development; (b) sex differences in degree of lateralization may not be found for all lateralized functions.

These findings have a number of interesting implications for theories of brain mechanisms underlying spatial and linguistic functions. A discussion of such implications in part III (Harshman and Berenbaum, 1975) includes the following possiblities: (a) there may be a causal relationship between sex differences in aspects of lateralization and sex differences in aspects of verbal and visuo-spatial performance; (b) differences in brain organization may also contribute to the female's relative immunity to certain language disorders; (c) existing knowledge about sex differences in performance can be used to suggest a number of useful lines of investigation of the role of lateralization in differential cognitive abilities, e.g. verbal fluency vs. comprehension, spatial orientation vs. visualization, etc; and (d) attempts to understand why men and women may have developed different biological trade-offs between brain organizations optimal for verbal vs. visuo-spatial functioning may lead to insights into the evolutionary role of cerebral lateralization in the development of human abilities.

References

Bakan, P. (1969) 'Hypnotizability, laterality of eye-movements and functional brain asymmetry.' *Perceptual and Motor Skills,* 28, 927-932.

Bakan, P. (1971) 'The eyes have it.' *Psychology Today*, April, 96: 64-67.

Bakan, P., and Putnam, W. (1974) 'Right-left discrimination and brain lateralization: Sex differences.' *Archives of Neurology* , 30: 334-335.

Bock, R.D. (1973) 'Word and image: Sources of the verbal and spatial factors in mental test scores.' *Psychometrika* , 38: 437-457

Bock, R.D. and Kolakowski, D. (1973) 'Further evidence of sex-linked major-gene influence on human spatial visualizing ability.' *American Journ. of Human Genetics,* 25: 1-14.

Bogen, J.E., DeZure, R., TenHouten, W.D., and Marsh, J.F. (1972) 'The other side of the brain IV: The A/P ratio .' *Bulletin of the Los Angeles Neurological Societies,* 37, 49-61.

Breckenridge, M.S., and Vincent, E.L. (1965) *Child Development* (5th ed.). Philadelphia: Saunders.

Broverman, D.M., and Klaiber, E.L. (1969) 'Negative relationships between abilities.' *Psychometrika*, 34: 5-20.

Brown, J.W., and Jaffe, J. (1975) 'Hypothesis on cerebral dominance.' *Neuropsychologia*, 13: 107-110.

Bryden, M.P. (1965) 'Tachistoscopic recognition, handedness, and cerebral dominance.' *Neuropsychologia*, 3: 1-8.

Buckley, F. (1971) 'Preliminary report on intelligence quotient scores of patients with Turner's syndrome: A replication study.' *British Journ. of Psychiatry*, 119: 513-514.

Buffery, A.W.H., and Gray, J.A. (1972) 'Sex differences in the development of spatial and linguistic skills." In C. Ounsted and D.C. Taylor (Eds.), *Gender differences: Their ontogeny and significance.* Edinburgh: Churchill Livingstone.

Corballis, M.C., and Beale, I.L. (1970) 'Bilateral symmetry and behavior.' *Psychological Review*, 77: 451-464.

Culver, C.M., Tanley, J.C., and Eason, R.G. (1970) 'Evoked cortical potentials: Relation to hand dominance and eye dominance.' *Perceptual and Motor Skills*, 30: 407-414.

Dawson, J.L.M. (1972) 'Effects of sex hormones on cognitive style in rats and men.' *Behavior Genetics*, 2: 21-42.

Day, M.E. (1967) 'An eye-movement indicator of type and level of anxiety: Some clinical observations.' *Journ. of Clinical Psychology*, 23: 433-441.

Duke, J.D. (1968) 'Lateral eye movement behavior.' *The Journ. of General Psychology*, 78: 189-195.

Dye, N.W., and Very, P.S. (1968) 'Growth changes in factorial structure by age and sex.' *Genetic Psychology Monographs* , 78: 55-88.

Eason, R.G., Groves, P., and Bonelli, L. (1967) 'Differences in occipital evoked potentials recorded simultaneously from both cerebral hemispheres in man.' *Proceedings of the 75th Annual Convention of the American Psychological Association*, 2: 95-96.

Eason, R.G., Groves, P., White, C.T., and Oden, D. (1967) 'Evoked cortical potentials: Relation to visual field and handedness.' *Science*, 156: 1643-1646.

Ehrlichman, H.I. (1971) Hemispheric functioning and individual differences in cognitive abilities (Doctoral dissertation, New School for Social Research). *Dissertation Abstracts International* , 1972, 33: 2319B. (University Microfilms No. 72-27, 869).

Etaugh, C.F. (1972) 'Personality correlates of lateral eye movement and handedness.' *Perceptual and Motor Skills* , 34: 751-754.

Garai, J.E., and Scheinfeld, A. (1968) 'Sex differences in mental and behavioral traits.' *Genetic Psychology Monographs* , 77: 169-299.

Gardner, R.W., Holzman, P.S., Klein, G.S., Linton, H.B., and Spence, D.P. (1959) 'Cognitive control: A study of individual consistencies in cognitive behavior.' *Psychological Issues*, 1: (Whole No. 4).

Garron, D.C., and Vander Stoep, L.R. (1969) 'Personality and intelligence in Turner's syndrome.' *Archives of General Psychiatry* , 21: 339-346.

Geschwind, N. (1974) The anatomical basis of hemispheric differentiation. In S.J. Dimond and J.G. Beaumont (Eds.), *Hemisphere function in the human brain*. New York: Wiley.

Gur, R.C., and Gur, R.E. (1974) 'Handedness, sex, and eyedness as moderating variables in the relation between hypnotic susceptibility and functional brain asymmetry. *Journ. of Abnormal Psychology* , 83: 635-643.

Harshman, R.A. and Berenbaum, S.A. (1975) Sex, language, and the brain, Part III: Interpretations and implications of adult sex differences in lateralization. Manuscript in preparation.

Harshman, R.A., and Krashen, S.D. (1974) 'On the development of lateralization. In *Cerebral Dominance* (Brain Information Service Conference Report No. 34). Los Angeles: UCLA Brain Information Service.

Harshman, R.A., Remington, R., and Krashen, S.D. (1975) Sex, language and the brain, Part II; Evidence from dichotic listening for adult sex differences in verbal lateralization. Manuscript submitted for publication.

Hartlage, L.C. (1970) 'Sex-linked inheritance of spatial ability.' *Perceptual and Motor Skills*, 31: 610.

Herron, J. 'EEG alpha asymmetry and dichotic listening in stutterers.' Paper presented at the Annual Convention of the Society for Psychophysiological Research, Salt Lake City, October 1974.

Johnson, P.O. (1973) Electrophysiological correlates of overt language behavior in humans. Unpublished master's thesis, California State University at San Francisco.

Kannan, P.M., and Lipscomb, D.M. (1974) 'Bilateral hearing asymmetry in a large population.' *Journ. of the Acoust. Soc. of America,* 55: 1092-1094.

Kimura, D. (1969) 'Spatial localization in left and right visual fields.' *Canadian Journ. of Psychology*, 23: 455-458.

Kimura, D. (1973) 'The asymmetry of the human brain.' *Scientific American,* 228:3, 70-78.

Kinsbourne, M. (1972) 'Eye and head turning indicates cerebral lateralization.' *Science*, 176: 539-541.

Kinsbourne, M. (1974) 'Direction of gaze and distribution of cerebral thought processes.'Neuropsychologia, 12, 279-281.

Krashen, S.D. (1972) Language and the left hemisphere. *Working Papers in Phonetics* 24: Univ. of Calif. Los Angeles. (University Microfilms No. 10,085).

Krashen, S., and Harshman, R. (1972a) 'Lateralization and the critical period.' *Journ. of the Acoust. Soc. of America,* 52: 174, (abstract).

Krashen, S., and Harshman, R. (1972b) 'Lateralization and the critical period.' *Working Papers in Phonetics* 23: Univ. of Calif. Los Angeles. 13-21. (University Microfilms No. 10,085).

Lansdell, H. (1961) 'The effect of neurosurgery on a test of proverbs.' *American Psychologist,* 16: 448, (abstract).

Lansdell, H. (1962) 'A sex difference in effect of temporal-lobe neurosurgery on design preference.' *Nature* , 194: 852-854.

Lansdell, H. (1968a) 'Effect of extent of temporal lobe ablations on two lateralized deficits.' *Physiology and Behavior* , 3: 271-273.

Lansdell, H. (1968b) 'The use of factor scores from the Wechsler-Bellevue Scale of Intelligence in assessing patients with temporal lobe removals.' *Cortex,* 4: 257-268.

Lansdell, H. (1973) 'Effect of neurosurgery on the ability to identify popular word associations.' *Journ. of Abnormal Psychology,* 81: 255-258.

Lansdell, H., and Davie, J.C. (1972) 'Massa intermedia: Possible relation to intelligence.' *Neuropsychologia* , 10: 207-210.

Lansdell, H., and Urbach, N. (1965) 'Sex differences in personality measures related to size and side of temporal lobe ablations.' *Proceedings of the 73rd Annual Convention of the American Psychological Association*, 113-114, (summary).

Levy, J. (1973) Lateral specialization of the human brain: Behavioral manifestations and possible evolutionary basis. In J. A. Kiger Jr. (Ed.), *The Biology of Behavior*. Corvallis, Oregon: Oregon State Univ. Press.

Maccoby, E. E., and Jacklin, C. N. (1974) *The Psychology of Sex Differences*. Stanford: Stanford Univ. Press.

Marcel, T., Katz, L., and Smith, M. (1974) 'Laterality and reading proficiency.' *Neuropsychologia* , 12: 131-139.

Marshal, J.C. and Holmes, J.M. (1974) 'Sex handedness, and differential hemispheric specialization for components of word perceptions.' *Int. Research Communication Service (Neurobiol., Psychol. Reprod.)* *2:* 1344.

Marshal, J.C., Newcombe, F. and Holmes, J.M. (1975) Lexical memory: A linguistic approach. In Kennedy, R.A. and Wilkes, A. (Eds) *Studies in long term memory*. New York: Wiley.

Masica, D.N., Money, J., Ehrhardt, A.A., and Lewis, V.G. (1969) 'IQ, fetal sex hormones and cognitive patterns: Studies in the testicular feminizing syndrome of androgen insensitivity.' *Johns Hopkins Medical Journ.* 124: 34-43.

McFarlane, M. (1925) 'A study of practical ability.' *British Journ. of Psychology Monographs*, No. 8.

McGlone, J., and Davidson, W. (1973) 'The relation between cerebral speech laterality and spatial ability with special reference to sex and hand preference.' *Neuropsychologia*, 11: 105-113.

McGlone, J., and Kertesz, A. (1973) 'Sex differences in cerebral processing of visuospatial tasks.' *Cortex* , 9: 313-320.

McKerracher, D.W. (1971) 'Psychological aspects of a sex chromatin abnormality.' *The Canadian Psychologist* , 12: 270-281.

Molfese, D.L. (1973) 'Cerebral asymmetry: The development of speech perception in the human brain.' Paper presented at Midwestern Psychological Assoc. Convention,Chicago, May.

Money, J. (1964) 'Two cytogenetic syndromes: Psychologic comparisons 1. Intelligence and specific-factor quotients.' *Journ. of Psychiatric Research*, 2: 223-231.

Mooney, C.M. (1957a) 'Age in the development of closure ability in children.' *Canadian Journ. of Psychology,* 11: 219-226.

Mooney, C.M. (1957b) 'Closure as affected by viewing time and multiple visual fixations.' *Canadian Journ. of Psychology,* 11: 21-28.

Mussen, P.H., Conger, J.J., and Kagan, J. (1963) *Child Development and Personality* (2nd ed.). New York: Harper and Row.

Newcombe, F. (1974) 'Selective deficits after focal cerebral injury. In S.J. Dimond and J.G. Beaumont (Eds.), *Hemisphere function in the human brain.* New York: Wiley.

Oetzel, R.M. (1966) Annotated bibliography and classified summary of research in sex differences. In E.E. Maccoby (Ed.), *The Development of Sex Differences.* Stanford University Press.

Ounsted, C. and Taylor, D.C. (1972) The Y chromosome message: A point of view. In C. Ounsted and D.C. Taylor (Eds.), *Gender differences: Their ontogeny and signficance.* Edinburgh: Churchill Livingstone.

Petersen, A.C. (1973) The relationship of androgenicity in males and females to spatial ability and fluent production. Unpublished doctoral dissertation, University of Chicago.

Porac, C. and Cohen, S. (1969) 'Is eye dominance a part of generalized laterality?' *Perceptual and Motor Skills*, 28: 927-932.

Scheinfeld, A. (1944) *Women and men.* New York: Harcourt, Brace and Co.

Sherman, J.A. (1974) 'Field articulation, sex, spatial visualization, dependency, practice, laterality of the brain and birth order.' *Perceptual and Motor Skills,* 38: 1223-1235.

Sinclair, D. (1969) *Human Growth after Birth.* London: Oxford Univ. Press.

Stafford, R.E. (1961) 'Sex differences in spatial visualization as evidence of sex-linked inheritance.' *Perceptual and Motor Skills,* 13: 428.

Tanner, J.M. (1970) Physical growth. In P.H. Mussen (Ed.), *Carmichael's manual of child psychology,* (3rd ed.) (Vol. 1). New York: Wiley.

Taylor, D.C. (1969) 'Differential rates of cerebral maturation between sexes and between hemispheres.' *Lancet, 2:* 140-142.

Terman, L.M., and Tyler, L.E. (1954) Psychological sex differences. In L. Carmichael (Ed.), *Manual of child psychology,* (2nd ed.). New York: Wiley.

Tyler, L.E. (1965) *The Psychology of Human Differences.* New York: Appleton-Century-Crofts.

Vandenberg, S.G. (Ed.) (1968) *Progress in Human Behavior Genetics.* Baltimore: Johns Hopkins Press.

Very, P.S., and Iacono, C.H. (1970) 'Differential factor structure of seventh grade students.' *The Journ. of Genetic Psychology*, 117: 239-251.

Wada, J.A. (1974) 'Morphologic asymmetry of human cerebral hemispheres: Temporal and frontal speech zones in 100 adult and 100 infant brains.' *Neurology*, 24: 349,(Abstract).

Weiten, W., and Etaugh, C. (1973) 'Lateral eye movement as related to verbal and perceptual-motor skills and values.' *Perceptual and Motor Skills*, 36: 423-428.

Weiten, W., and Etaugh, C. (1974) 'Lateral eye-movement as a function of cognitive mode, question sequence, and sex of subject.' *Perceptual and Motor Skills*, 38: 439-444.

Witelson, S.F., and Pallie, W. (1973) 'Left hemisphere specialization for language in the newborn: Neuroanatomical evidence of asymmetry.' *Brain*, 96, 641-646.

Wolf, S.M. (1973) 'Difficulties in right-left discrimination in a normal population.' *Archives of Neurology*, 29: 128-129.

Yen, W.M. (1975) 'Sex-linked major-gene influences on selected types of spatial ability.' *Behavior Genetics*, 5, 281-298.

Footnotes

[1] This article is based, in part, on a paper presented to the UCLA Conference on Human Brain Function, September, 1974.

[2] The authors would like to express their deep appreciation to Peter Ladefoged, Director of the UCLA Phonetics Laboratory, for the continuing support which made this study possible. We are indebted to Sheri Berenbaum, Joseph E. Bogen, Elizabeth Hecht and Stephen Krashen for important suggestions and invaluable assistance, and to Allan and Louise for both personal encouragement and financial help. We would also like to thank Lissy F. Jarvik for providing us with her unpublished manuscript on sex differences in intelligence. Others who read earlier versions of this article and provided helpful comments or contributions include: Suzanna deRooy, Louis Goldstein, Steven Greenberg, George Papçun, and William S-Y. Wang. Although we thank them for their contributions, any errors which we have perversely insisted on including remain our sole responsibility.

[3] Roger Remington is now at the Department of Psychology, University of Oregon.

[4] Among patients with right hemisphere damage, both sexes showed about equal impairment of performance on a spatial (closure) task. However, among left damaged patients, males showed performance in the normal range (when compared to data reported by Mooney, 1957a, 1957b; Newcombe, 1974, p. 318) while females showed performance which was *superior* to that normally found in healthy subjects. Thus, while the difference in spatial performance between right and left hemisphere groups was greater for the females, this fact was due to an anomolous superiority of a group of brain damaged subjects. It is not clear whether these data suggest greater female lateralization of this spatial task, unusual patterns of hemisphere involvement on this task, or simply indicate an atypical sample of left damaged female patients.

[5] Other sex differences related to lateral eye movements have also been noted. It has been reported by Duke (1968) and informally by Bakan (1971) that females show less consistent eye movements, but this has not been found by Weiten and Etaugh (1974). On the other hand, females were found to show fewer right and more left movements according to Day (1967), Weiten and Etaugh (1974) and Marshalek (personal communication, 1975), but this was not found by Etaugh (1972). Such variation in experimental results is not surprising, since patterns of reflective eye movements appear to be strongly affected by differences in experimental procedure which appear to bring out different aspects of the phenomenon (see Kinsbourne, 1974). With the exception of Gur and Gur, however, the sex differences in eye movements reported to date seem ambiguous: they might simply reflect culturally induced sex differences in preferred modes of problem solving, rather than differences in cerebral lateralization of a given mode of problem solving (for further discussion on distinguishing between these two types of effects, see Harshman, 1975). Perhaps both physiological and cultural effects are involved. Hopefully, future eye movement research aimed specifically at sex differences will try to disentangle these various factors.

The Interface between Phonetics and Phonology

Victoria A. Fromkin

(Paper presented at the 8th International Congress
of Phonetic Sciences, August 17-23, 1975
Leeds, England)

It is seventy three years since the great phonetician, Henry Sweet, wrote: "My own subject, Phonetics, is one which is useless by itself, while at the same time it is the foundation of all study of language whether theoretical or practical." (Firth, 1957)

I think it is in keeping with the spirit of Sweet's statement to go even further and say that the study of phonetics is impossible by itself, since an understanding of language in general and of specific linguistic systems in particular is prerequisite to even determining what is or is not a discrete speech sound or unit, a primitive in phonetics. In addition, historically, phonology has raised the questions which have been most influential in determining the kinds of phonetic research undertaken.

But it is equally clear that the study of phonology cannot be divorced from phonetics if phonologists aim to provide explanatory and predictive theories of sound systems, rather than merely descriptive theories.

The traditional question of whether phonetics should be considered a discipline separate from phonology is thus a meaningless one. The strict separation between phonetics as the study of 'parole' from phonology, the study of 'langue', as first proposed by Saussure, and then reiterated by Trubetzkoy, if maintained today, would impede research in both areas. This is not to negate the statement of Trubetzkoy that each discipline use "quite different methods for investigation: the study of sounds of the speech event which is concerned with concrete physical phenomena must use the methods of natural science, the study of sounds of the system of language...purely linguistic methods..." (Trubetzkoy 1939) Trubetzkoy, himself, never divorced language from its phonetic base, pointing out that "...speech events and systems of language presuppose each other. They are inseparably linked to each other and can be considred as two interrelated aspects of the same phenomenon, 'language'."

The importance of this view has become more and more apparent as work in both phonetics and phonology has developed. There are posited phonological universals concerned with possible sound systems, 'unmarked' (natural) versus 'marked' (unnatural) rules, phonetic features, historical sound changes, to name just a few, which can only be understood, verified, and explained by reference to the physiological and acoustic aspects of speech.

While this is now, hopefully, the accepted view, I think it important to reaffirm the interface between phonetics and phonology. To do this, in this short paper, I shall refer to a question which has interested phoneticians far more than it has phonologists -- the universally attested duration distinction of vowels before voiced and voiceless obstruents. Phonologists, recognizing this phenomenon, have treated it in one of two ways -- either by positing a universal vowel lengthening rule which would apply to all languages, or by including such a rule in the particular grammars of individual languages. Clearly neither approach provides an explanation. In fact, without phonetic data one cannot decide between the two proposals -- universal rule versus language specific rules. And, as will be discussed below, different implications may follow from the two 'solutions'.

One posited phonetic explanation for the phenomenon is that the increased duration is due to the greater rate of articulation movement in the production of voiceless stops. Since this is language independent one might conclude that no language specific rules are required. But as Chen (1969) and others have shown, there are languages, such as English, in which the increased duration exceeds the evidenced 27 ms. resulting from this physiological or mechanical cause. Since the duration distinction in English is not predictable from solely automatic causes, there must be a specific rule in the grammar. This rule represents the fact that in learning English, speakers must be aware of the allophonic length distinction, in order to produce it. Thus, a universal rule is necessary but not sufficient. Cross linguistic experimental data, and data relating to the timing of physiological mechanisms are required by the phonologists in deciding whether a language has such a rule in its grammar.

To understand how such a rule may arise, perceptual as well as physiological studies are required. Fortunately, there is much phonetic research of both kinds. (cf. Denes 1955; Lisker 1970, 1974; Lehiste 1970; Massaro 1974; Nooteboom 1972, 1974, et al.)

It has, for example, been shown that "English speaking listeners use the differences in vowel length as a perceptual cue to the voicing of the following consonant." (MacNeilage 1972) This is further evidence for the fact that despite the fact that vowel duration is a redundant feature in English, such length distinctions must be preprogrammed or in the grammar. Without such kinds of evidence, the rules posited by phonologists remain unverified, even if true.

Phonetic data of this kind also has immediate relevance to the processes of historical change. It would not be surprising if languages such as English would begin to use vowel length duration phonemically with a neutralizaiton of the voiced/voiceless contrast of obstruents in final position arising. Phonetic research predicts such a possible change. It may already be underway in English with the devoicing or partial devoicing of final consonants.

We are only beginning to understand how different acoustic cues -- redundant and distinctive -- are used by speakers in the speech understanding process. What we learn will or should play a profound role in the development of phonological theory, such as the set of phonetic features necessary to classify phonemic segments. For example, a preliminary study by Louis Goldstein of UCLA suggests that when all other cues as to the voiced/voiceless contrast of final consonants are held constant, e.g. vowel duration and voicing, pitch alone may be sufficient to distinguish such minimal pairs as <u>cat</u> and <u>cad</u>. If the extended study underway reaffirms the finding that subjects will consistently identify those stimuli with a higher F_0 as ending with a voiceless stop, this may contribute to our understanding of the genesis of tone. Furthermore, such a result will show that speakers are aware -- unconsciously to be sure -- of aspects of the acoustic signal which are not usually specified in a phonological representation but which clearly constitutes part of their linguistic knowledge.

In may be inappropriate or foolish to speak of the importance of phonetic research at this Congress of Phonetic Sciences, a subject more suitable to be discussed with theoretical phonologists. But I think the interface between phonetics and phonology is worthy of discussion and is a question of concern to us all.

References:

Chen, M. (1969) 'Vowel length variation as a function of the voicing of the consonant environment' *Project on Linguistic Analysis*, #9. University of California, Berkeley, Phonology Laboratory.

Denes, P. B. (1955) 'The effect of duration on the perception of voicing' *JASA* 27:761-764.

Firth, J. R. (1957) *Papers in Linguistics 1934-1951* London: Oxford University Press.

Lehiste, I. (1970) *Suprasegmentals* Cambridge, Mass.: MIT Press.

Lisker, L. (1970) 'On time and timing in speech' Haskins Laboratories *Status Report on Speech Research*

Lisker, L. (1974) 'On explaining vowel duration variation' *Glossa* 8.2:233-246.

MacNeilage, P. (1972) 'Speech physiology' in *Proceedings of Symposium on Speech Production-Speech Perception: Their Relationship to Cortical Functioning* Edited by J. Gilbert. Academic Press.

Massaro, D. (1974) 'Perceptual units in speech recognition' *J. Exper. Psych.* 102.2:199-208.

Nooteboom, S. G. (1972) 'Some timing factors in the production and perception of vowels' *Occasional Papers*, 13. University of Essex.

Nooteboom, S. G. (1974) 'Some context effects on phonemic categorization of vowel duration' *IPO Annual Progress Report*, #9:47-55.

Trubetzkoy, N. (1939) *Principles of Phonology* (C. Baltaxe, trans. 1969) Berkeley: Univ. of California Press.

Two "Voiceprint" Cases

George Papçun and Peter Ladefoged

[Paper presented at the 86th Annual Meeting for the Acoustical
Society of America, October 30 - November 2, 1973,
Los Angeles, California]

At about 9:15 p.m. in Riverside, California on April 2, 1971, the police received what appeared to be a routine call reporting a house burglary. That telephone call, like all incoming calls, was recorded on a 24 hour continually recording, multiple channel recorder. The transcript of the call follows:

Police: Riverside Police

Unknown: Yes, uh, my name is Keith Davis, and I live at 4792 Ottawa.

Police: Yessir

Unknown: And I wanna report a house burglar.

Police: O. K. Did this just happen or did you just find it?

Unknown: Uh, I don't know. I was away 2 - 3 hours, and when I just come home I found my st . . .

Police: O. K., and that's 4792 Ottawa?

Unknown: Yes

Police: O. K. Now it may be a while. We're gonna . . . there as soon as we can.

Unknown: O. K. . . . Uh . . .

Police: It'll probably be about 10 - 15 minutes.

Unknown: O. K.

Police: O. K. Thank you. 'M bye.

I WAS A - WAY

Figure 1. Spectrogram of unknown call in Jackson case. Note frequency limitations and telephone ringing which combines with and obscures the pattern of the second formant.

In response to the call, two officers were dispatched to 47 92 Ottawa. Some ten minutes later, over the police radio, came the call "Officer needs help!" When other officers arrived at the scene, the two officers were found shot and killed.

In Figure 1 we see the extreme limitation in the frequency range of the recording of the incoming call introduced by the fact that the 24 hour recorder had a tape travel speed of 15/16 i.p.s. Useful information does not extend over about 2300 Hz.

In Figure 1 we can also see examples of the background noise which was introduced at the mouthpiece of the telephone of the officer who received the incoming call because he was seated in a busy office near the police dispatcher. In the horizontal bar representing the ring of a telephone, you can discern the individual "dings" of the bell. In the region of the second formant in the words "I was," you can see that the pattern of the bell combines with and obscures the pattern of the second formant. It is possible to wire telephones with suppression circuits so that external noise would not be introduced into the recording.

Figure 2 shows spectrograms of the phrase "I live at 4792 Ottawa," wideband top, narrowband, bottom. From the narrowband spectrogram, notice that the noise continues even in places where it is not clearly evident on the wideband spectrogram.

Also, notice some remarkable anomalies: between the first and second syllables of the word "seven" there is a *break* downward in pitch, and in the second formant region there is a confused harmonic pattern indicating, perhaps, that more than one voice source is being analyzed. The words "92," indicated by the bottom marks, are on a *very* low pitch as compared with the rest of the call. In spite of these anomalies, the word "nine" was used as a point of comparison between the unknown and the suspect.

The points mentioned raise the issue of how to extrapolate from experiments in which frequency limitations, background noises, and such anomalies as reported here, were not studied, to forensic situations in which they *do* occur.

The case of the People vs. Chapter involves some of the same considerations as well as some additional factors. In this case Steven Chapter was accused of having made a bomb threat to the telephone company using a telephone line that was normally available only to telephone repairmen. The group of telephone repairmen who normally used that telephone line were all asked to record the words in the bomb threat. All of them complied with the request except Mr. Chapter who refused on the grounds that it violated his personal freedom. His voice was surreptitiously recorded by telephone company investigators when he called in a work order. This recording, together with recordings of people who had voluntarily recorded the exact words of the bomb threat, were assigned numbers and sent to the Voice Identification Unit of the Michigan State Police Force. The surreptitiously obtained recording of Steven Chapter was of course very different

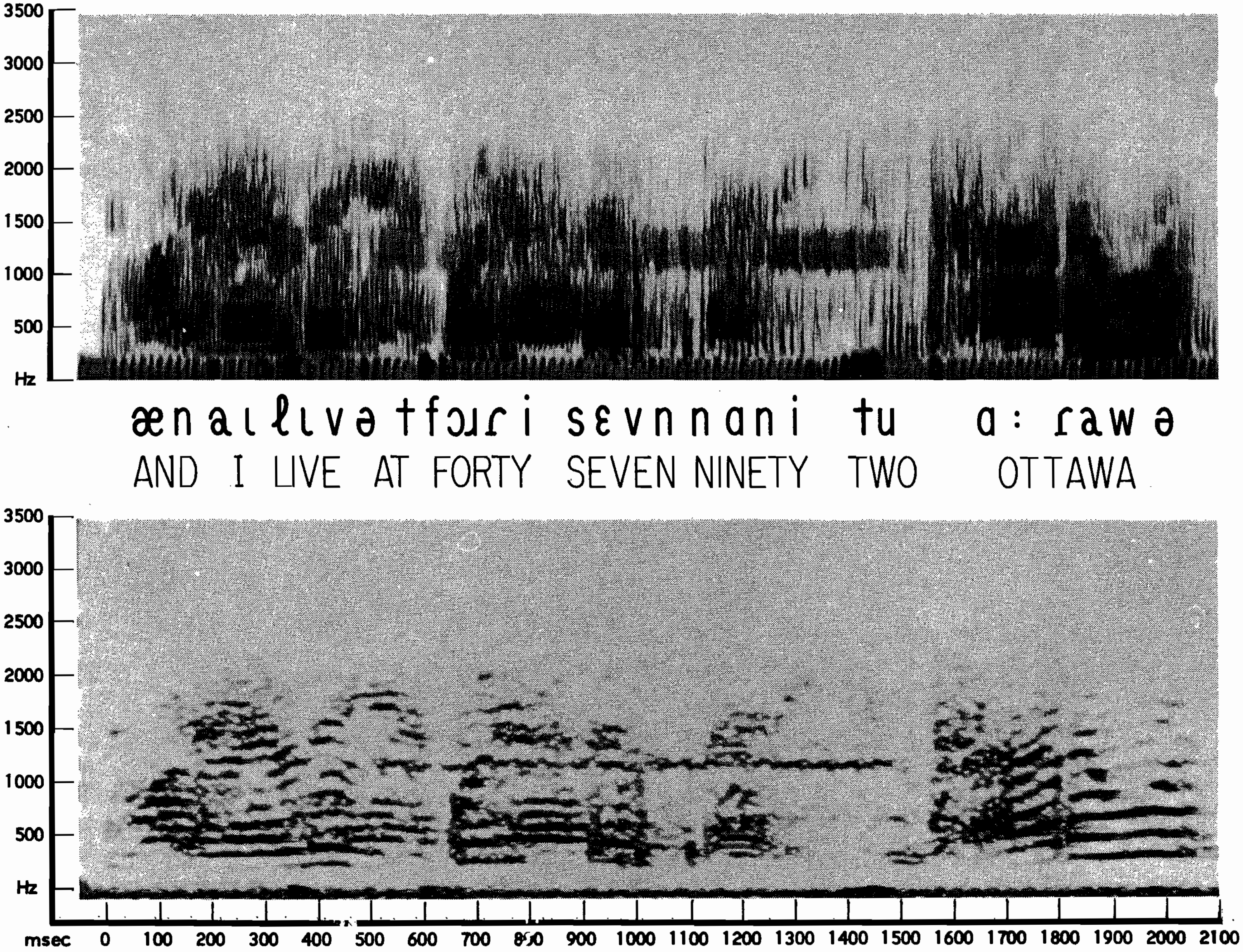


Figure 2. Spectrograms of unknown call in Jackson case; broadband top, narrowband bottom. Note frequency limitations and the break downward in pitch in the words "ninety-two".

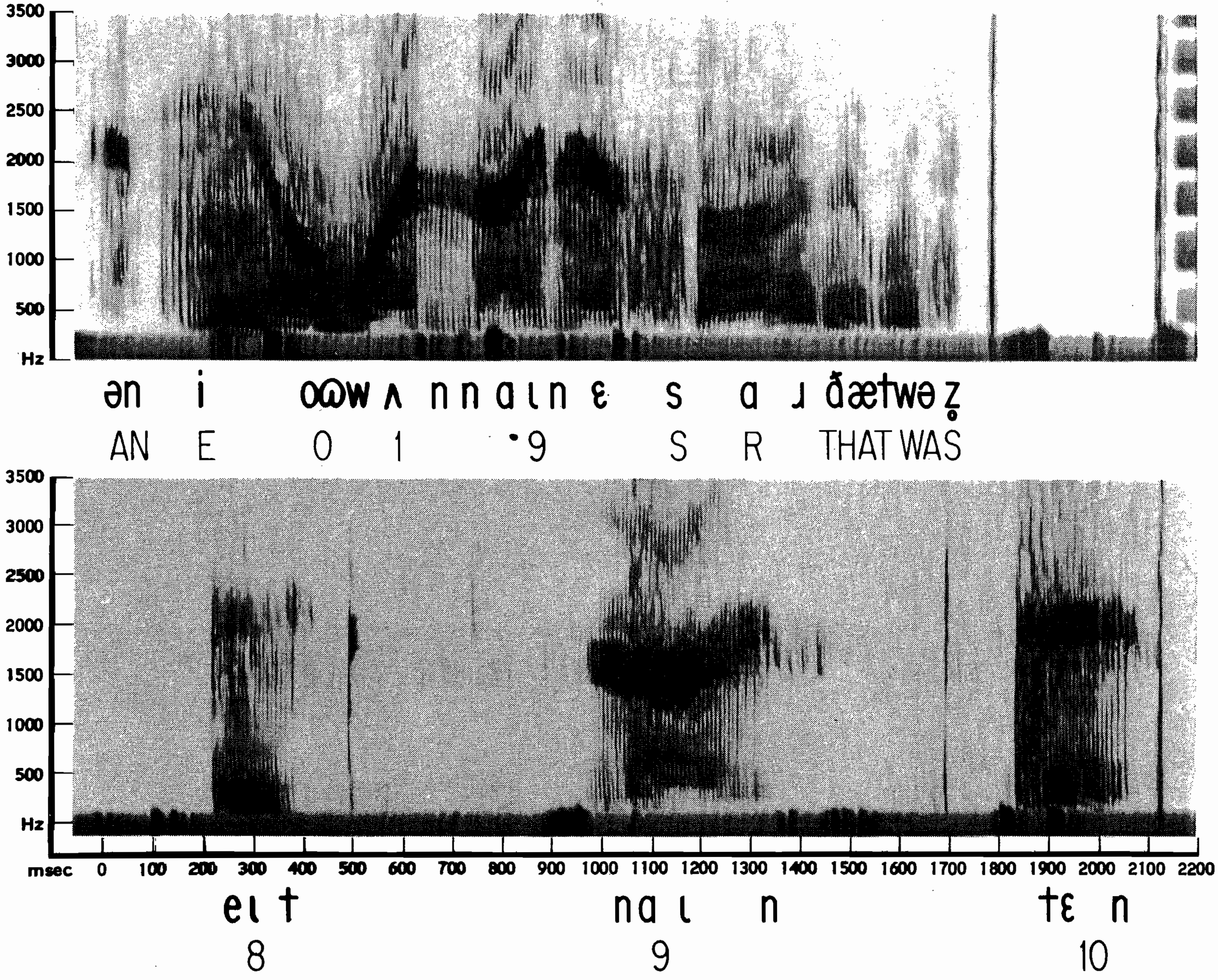


Figure 3. Top: Spectrogram of unknown call in Chapter case. Bottom: Surreptitiously recorded voice of Stephen Chapter. The number "eight" was matched with the letter "E".

in quality as well as in content from the other recordings. From his examination, Lt. Nash of the Voice Identification Unit of the Michigan State Police Force concluded that Mr. Chapter's voice was the voice of the unknown caller. Lt. Nash's evidence to this effect was used in the preliminary legal proceedings. Before the actual trial the District Attorney obtained a court order requiring Mr. Chapter to say the same words as on the bomb threat and he sent this new recording, together with a number of recordings of people in his office and local peace officers all saying the same words as the bomb threat call, to Lt. Nash. What happened next is not quite clear. According to the District Attorney, who was called as a witness by the defense, in a telephone conversation Lt. Nash tentatively identified a deputy District Attorney as the voice of the unknown caller. According to Lt. Nash, in his official report he identified Mr. Chapter as the unknown caller. This report was made after the District Attorney had told Lt. Nash which of the numbers corresponded to Mr. Chapter.

In this case, as in the other case, the frequency response of the recordings was limited. As you can see from Figure 3, the highest frequencies visible are 2500 Hz. This slide also presents another problem that occurred in this case. The phrase shown, "an E 0 1 9 S R," was taken from the surreptitiously obtained recording of Mr. Chapter's voice calling in a work order to his office. We know that these were the words used because a copy of the work order from which Mr. Chapter was reading was presented in evidence. Nevertheless Lt. Nash claimed that the phrase was "an A 0 1 9 S R." This is an easy mistake to make considering the fact that the whole phrase is simply telephoneman's jargon, and Mr. Chapter's accent is such that he pronounces the name of the fifth letter of alphabet, "e," very much like many other people pronounce the first letter of the alphabet, "a." But this word, which we know to be "e," was compared with the number "eight" as said by the unknown voice as shown in the lower part of the figure. You can see from this figure that the frequency response of the recording of the unknown voice was limited so that there is nothing much over 2500 Hz available. In addition to this faulty comparison there were other mistakes in the labeling of the spectrograms. We know from listening to the recording and from reading the work order what Mr. Chapter must have said. Thus in one case he said "8 2 2 T," but this phrase was labeled "8 2 2 2." Nevertheless it was on the basis of evidence such as this that Lt. Nash declared under oath that "the voice of the unknown is that of Stephen Chapter and could be that of no other person in the world." These two cases raise the issue of how to extrapolate from laboratory studies to forensic situations.

We would like to put on record that we both believe that it is possible to use spectrographic data to gain evidence about the degree of similarity of two voices. We further believe that when a scientist has carefully compared appropriate recordings, his opinion of the degree of similarity between the two voices has evidential value. The degree of reliability of an identification based on voice similarity will vary according to many circumstances which must be ascertained and considered by competent scientists. The weight such evidence should be considered to have in particular cases is a matter for the courts to decide in view of all the evidence.

Index of publications by members of the UCLA Phonetics Laboratory 1971 -- September 1975

Compiled by Sue Gass

George D. Allen

1971 'Acoustic level and vocal effort as cues for the loudness of speech' *Working Papers in Phonetics* 17:29-57

'Syllable structure and sentence rhythm' *Working Papers in Phonetics* 19:55-60

--, and Peter Ladefoged. 'Syllable structure and sentence rhythm - a cross language study' Abstracted in *Journal of the Acoustical Society of America* 50:116. Also published in Rigault, Andre and Rene Charbonneau (eds.) *Proceedings of the VII International Congress of Phonetic Sciences* pp. 833-836. The Hauge: Mouton.

Kay Atkinson-King

1973 'Children's acquisition of phonological stress contrasts' *Working Papers in Phonetics* 25.

Lorian Baker

1974 'The lexicon: some psycholinguistic evidence' *Working Papers in Phonetics* 26.

Susan Curtiss

1973 --, Stephen Krashen, Victoria Fromkin, David Rigler and Marilyn Rigler. 'Language acquisition after the critical period: Genie as of April 1973' *Papers from the 9th Regional Meeting of the Chicago Linguistics Society* pp. 98-103.

(See Stephen Krashen, 'Language lateralization in a case of extreme psychosocial deprivation')

1974 --, Victoria Fromkin, Stephen Krashen, David Rigler, and Marilyn Rigler. 'The linguistic development of Genie' *Language* 50:528-554

(See Victoria Fromkin, 'The development of language in Genie: a case of language acquisition beyond the "Critical Period"'.)

Baruch Elimelech

1974 'On the reality of underlying contour tones' *Working Papers in Phonetics* 27:74-83.

'Tone alternations in the Etsako verb' *Working Papers in Phonetics* 27:63-73.

Victoria A. Fromkin

1971 'The hunting of the 'quark' in speech production: further evidence from speech errors' *Journal of the Acoustical Society of America* 50:115-116.

'In defence of systematic phonemics' *Journal of Linguistics* 7:75-83.

'The non-anomalous nature of anomalous utterances' *Language* 47:27-51.

'Simplicity is a complicated question' *Working Papers in Phonetics* 17:76-84. Abstracted in *Working Papers in Phonetics* 19:61.

'There's many a slip twixt the brain and the lip' *Working Papers in Phonetics* 17:3. Abstracted in *Journal of the Acoustical Society of America* 49:94.

'Tone features and tone rules' *Working Papers in Phonetics* 21:50-57. Also published in *Studies in African Linguistics* 3:1 (1972).

'When is a non-occuring 'possible' verb impossible?' *Working Papers in Phonetics* 17:4-5.

1972 'Discussion paper on speech physiology' *Working Papers in Phonetics* 23:37-60. Also published in *Proceedings of Symposium on speech production and speech perception: their relationship to cortical functioning* Academic Press

'On the reality of linguistic constructs. Evidence from speech errors' *Working Papers in Phonetics* 22:103-110. Also published in *Proceedings of the VII International Congress of Phonetic Sciences* pp. 1107-1110. The Hauge: Mouton.

'Tone features and tone rules' *Studies in African Linguistics* 3:1.

1973 'Comments on "some unsolved problems in generative phonology" by R. N. Srivastava. *International Journal of Dravidian Linguistics* 2:2.

'Review of E. J. A. Henderson (ed.) *The Indispensable Foundation: a selection from the writings of Henry Sweet*' *General Linguistics* 13:1.

'Slips of the tongue' *Scientific American* December:111-117

Speech errors as linguistic evidence The Hague: Mouton.

(See Susan Curtiss, 'Language acquisition after the critical period: Genie as of April 1973'.)

(See Stephen Krashen, 'Language lateralization in a case of extreme psychosocial deprivation'.)

(See Diana Van Lancker, '"Tone" and pitch perception'.)

1974 'On the phonological representation of tone' *Working Papers in Phonetics* 27:1-17.

'Review of *Hesitation phenomena in adult aphasic and normal speech* by Gerd Quinting' *Lingua*

'Explaining the explanation' *Georgetown University Round Table on Language and Linguistics* pp. 1-13. Georgetown University Press.

--, Stephen Krashen, Susan Curtiss, David Rigler and Marilyn Rigler. 'The development of language in Genie: a case of language acquisition beyond the "Critical Period"' *Brain and Language* 1:81-107.

--, and R. Rodman. *An Introduction to Language* New York: Holt, Rinehart and Winston.

(See Susan Curtiss, 'The linguistic development of Genie'.)

1975 -- and others. W. Lehmann (ed.) *Language and linguistics in the People's Republic of China* Austin: University of Texas Press

Jack Gandour

1974 'The glottal stop in Siamese: predictability in phonological description' *Working Papers in Phonetics* 27:84-91. Also published in *Phasaa* 4.2:66-78.

'Consonant types and tone in Siamese' *Working Papers in Phonetics* 27:92-117. Also published in *Journal of Phonetics* 2.4:337-350.

'An annotated bibliography on tone' *Working Papers in Phonetics 28*, University of California at Los Angeles, Compiled with Ian Maddieson and others.

'On the representation of tone in Siamese' *Working Papers in Phonetics* 27:118-146. Also published in Harris, Jimmy G. and James R. Chamberlain (eds.) *Studies in Tai Linguistics in Honor of William J. Gedney* pp. 170-195. Central Institute of English Language: Bangkok, Thailand.

'The features of the larynx: n-ary or binary?' *Working Papers in Phonetics* 27:147-159. Also published in *Phonetica* 32.4: 241-253.

Louis Goldstein

1974 --, and James R. Lackner. 'A sideways look at dichotic listening' *Journal of the Acoustical Society of America* 55:S10.

(See Richard Harshman, 'Factors underlying the articulatory and acoustic structure of vowels'.)

James R. Lackner and Louis Goldstein. 'Primary auditory stream segregation of repeated consonant-vowel sequences' Abstracted in *Journal of the Acoustical Society of America* 56:165.

1975 --, and Jamres R. Lackner. 'Alterations of the phonetic coding of speech sounds during repetition' *Cognition* 2:279-297.

James R. Lackner and Louis M. Goldstein. 'The psychological representation of speech sounds' *Quarterly Journal of Experimental Psychology* 27:173-185.

Richard Harshman

1971 'PARAFAC: An "explanatory" factor analysis procedure' *Working Papers in Phonetics* 19:14-16. Abstracted in *Journal of the Acoustical Society of America* 50:117.

(See Peter Ladefoged, 'Factor analyses of tongue shapes'.)

(See Peter Ladefoged, 'Parameters of tongue shapes'.)

(See Mona Lindau, 'Factor analyses of formant frequencies of vowels'.)

(See Dale Terbeek, 'Cross-language differences in the perception of natural vowel sounds'.)

1972 'Determination and proof of minimum uniqueness conditions for PARAFAC 1' *Working Papers in Phonetics* 22:111-117.

'PARAFAC 2: Extensions of a procedure for "explanatory" factor-analysis and multi-dimensional scaling' *Working Papers in Phonetics* 22:7. Abstracted in *Journal of the Acoustical Society of America* 51:111.

'PARAFAC 2: Mathematical and technical notes' *Working Papers in Phonetics* 22:30-44.

'PARAFAC: Parallel-factor models yielding mathematically unique, empirically meaningful solutions with certain classes of three-mode data' *Working Papers in Phonetics* 22:8-9.

--, and Stephen Krashen. 'An "unbiased" procedure for comparing degree of lateralization of dichotically presented stimuli' *Working Papers in Phonetics* 23:3-12. Abstracted in *Journal of the Acoustical Society of America* 52:174 and *Working Papers in Phonetics* 22:5.

(See Stephen Krashen, 'Lateralization and the critical period'.)

(See Peter Ladefoged, 'Control of the tongue in vowels'.)

(See Dale Terbeek, 'Is vowel perception non-Euclidean?'.)

1974 'Meaning and measurement of "degree of lateralization"' *Journal of the Acoustical Society of America* 55:S10.

--, Peter Ladefoged, Louis Goldstein and Joseph De Clerk. 'Factors underlying the articulatory and acoustic structure of vowels' Abstracted in *Journal of the Acoustical Society of America* 55:385.

--, and Stephen Krashen. 'On the development of lateralization' In *Cerebral Dominance* (Brain Information Service Conference Report 34) Los Angeles: UCLA Brain Information Service.

(See Geroge Papcun, 'Is the left hemisphere specialized for speech, language, and/or something else?'.)

(See Roger Remington, 'A possible sex difference in degree of lateralization of dichotic stimuli?'.)

Leon Jacobson

1971 'Index of publications by members of the UCLA Phonetics Laboratory 1963-1970' *Working Papers in Phonetics* 17: 104-116.

1972 (See Mona Lindau, 'The feature advanced tongue root'.)

Stephen Krashen

1971 (See George Papcun, 'Is the left hemisphere specialized for speech, language or something else?')

1972 'Language and the left hemisphere' *Working Papers in Phonetics* 24.

--, and Richard Harshman. 'Lateralization and the critical period' *Working Papers in Phonetics* 23:13-21. Abstracted in *Journal of the Acoustical Society of America* 52:174 and *Working Papers in Phonetics* 22:6.

(See Richard Harshman, 'An "unbiased" procedure for comparing degree of lateralization of dichotically presented stimuli'.)

(See George Papcun, 'Is the left hemisphere specialized for speech, language or something else?'.)

1973 --, Victoria Fromkin, Susan Curtiss, David Rigler and Sarah Spitz. 'Language lateralization in a case of extreme psychosocial deprivation' Abstracted in *Journal of the Acoustical Society of America* 53:367.

(See Susan Curtiss, 'Language acquisition after the critical period: Genie as of April 1973'.)

(See Susan Curtiss, 'The linguistic development of Genie')

(See Victoria Fromkin, 'The development of language in Genie: a case of language acquisition beyond the "Critical Period"'.)

Peter Ladefoged

1971 'An alternative set of vowel shift rules' *Working Papers in Phonetics* 17:25-28.

'A basic linguistic phonetics laboratory' *Working Papers in Phonetics* 17:25-28.

'The limits of phonology' *Form and substance:* papers presented to Eli-Fischer Jorgensen, Copenhagen: Akademisk Forlag, pp. 47-56.

'An opinion on "voiceprints"' *Working Papers in Phonetics* 19:84-87.

'Phonetics' *Working Papers in Phonetics* 20. Published in 15th Edition of the *Encyclopaedia Brittanica* 275-283 (1974).

'Phonological features and their phonetic correlates' *Working Papers in Phonetics* 21:3-12. Also published in *Journal of the International Phonetic Association* 2:2-12 (1972).

Preliminaries to linguistic phonetics Chicago: University of Chicago Press.

'The features of the larynx' *Journal of Phonetics* 1.1:73-83.

--, J.F.K. Anthony and Cordell Riley. 'Direct measurement of the vocal tract' *Working Papers in Phonetics* 19:4-13. Abstracted in *Journal of the Acoustical Society of America* 49:104.

--, Joseph L. De Clerk and Richard Harshman. 'Parameters of tongue shape' *Working Papers in Phonetics* 19:47.

--, Joseph L. De Clerk, and Richard Harshman. 'Factor analyses of tongue shapes' *Working Papers in Phonetics* 19:47. Abstracted in *Journal of the Acoustical Society of America* 50:114.

--, Ruth Glick and Clive Criper. *Language in Uganda* Nairobi, Kenya: Oxford University Press.

--, Raymond Silverstein and George Papcun. 'The interruptibility of speech' *Working Papers in Phonetics* 17:85-94. Published in *Journal of the Acoustical Society of America* 54:1105-1108 (1973).

(See George D. Allen, 'Syllable structure and sentence rhythm-a cross language study'.)

Clive Criper and Peter Ladefoged. 'Linguistic complexity in Uganda' In W. H. Whiteley, (ed.) *Language use and Social Change* London: Oxford University Press pp. 145-159.

(See Mona Lindau, 'Factor analyses of formant frequencies of vowels'.)

(See Theo Vennemann, 'Phonetic features and phonological features'.)

1972 'Phonetic prerequisites for a distinctive feature' In A. Waldman (ed.) *Papers in linguistics and phonetics in memory of Pierre Delattre* The Hague: Mouton.

'Phonological features and their phonetic correlates' *Journal of the International Phonetic Association* 2:2-12.

'The three glottal features' *Working Papers in Phonetics* 22:95-101. Abstracted in *Journal of the Acoustical Society* 51:101.

--, Joseph De Clerk, Mona Lindau and George Papcun. 'An auditory-motor theory of speech production' *Working Papers in Phonetics* 22:48-75.

--, Joseph De Clerk,and Richard Harshman. 'Control of the tongue in vowels' In Andre Rigault and Rene Charbonneau (eds.) *Proceedings of the Seventh International Congress of Phonetic Sciences* The Hague: Mouton pp. 349-354.

(See Mona Lindau, 'The feature advanced tongue root'.)

(See Ralph Vanderslice, 'Binary suprasegmental features and transformational word-accentuation rules'.)

1973 --, Raymond Silverstein and George Papcun. 'Interruptibility of speech' *Journal of the Acoustical Society of America* 54:1105-1108.

(See Theo Vennemann. 'Phonetic features and phonological features'.)

1974 'Phonetics' 15th Edition of *Encyclopaedia Brittanica* pp. 275-283.

'Respiration, laryngeal activity and linguistics' In Barry Wyke (ed.) *Ventilatory and Phonatory Control Systems* London: Oxford University Press pp. 299-314.

--, Iris Kameny and William A. Brackenridge. 'Acoustic effects of style of speech' *Journal of the Acoustical Society of America* 56:S28.

(See George Papcun, 'Two "voiceprint" cases'.)

1975 *A Course in Phonetics* New York: Harcourt Brace and Jovanovich.

Carl R. La Velle

1974 'An experimental study of Yoruba tone' *Working Papers in Phonetics* 27:160-170.

Mona Lindau

1971 --, Richard Harshman and Peter Ladefoged. 'Factor analyses of formant frequencies of vowels' *Working Papers in Phonetics* 19:17-25. Abstracted in *Journal of the Acoustical Society of America* 50:117.

1972 --, Leon Jacobson and Peter Ladefoged. 'The feature advanced tongue root' *Working Papers in Phonetics* 22:76-94. Abstracted as 'Cross linguistic study of advanced tongue root' *Journal of the Acoustical Society of America* 51:101.

1974 'Different vocal tract shapes with similar formant frequencies: evidence from Twi' *Journal of the Acoustical Society of America* 55:385.

1975 'Features for vowels' *Working Papers in Phonetics* 30.

Ian Maddieson

1974 'A note on tone and consonants' *Working Papers in Phonetics* 27:18-27.

'A possible new cause of tone-splitting evidence from Cama, Yoruba and other languages' *Working Papers in Phonetics* 27:28-46.

'The tone tome' (ed.) *Working Papers in Phonetics* 27.

'An annotated bibliography on tone' (ed.) *Working Papers in Phonetics* 28.

1975 --, and Kay Williamson. 'Jarawan Bantu' *African Languages/Langues Africaines* 1.1:125-163.

'Tone sandhi and intrinsic pitch of vowels in Foochow' In M. Noonan (ed.) *Papers from the 5th California Linguistic Association Conference* San Jose State University.

Laura Meyers

1972 'Polycategorial lexical attachment: a look at the translational lexicon proposed by Jeffrey Gruber' In Schachter and Bedell (eds.) *Critiques of Syntactic Studies* vol. II.

'Sai and Har: only, even and until in Hausa' *Proceedings of the 3rd Annual Conference on African Linguistics* Bloomington, Indiana.

1974 'The Hausa particle ko: an uncertainty modality' *Studies in African Linguistics* 5.3.

'Tone patterns in Hausa: a reanalysis of Hausa downdrift' *Working Papers in Phonetics* 27:47-62.

Sven Ohman

1974 (See Lloyd Rice, 'On the relationship between formant bandwidths and vocal-tract shape features'.)

George Papcun

1971 --, Stephen Krashen and Dale Terbeek. 'Is the left hemisphere specialized for speech, language or something else?' *Working Papers in Phonetics* 19:69-77. Abstracted in *Journal of the Acoustical Society of America* 51:79 (1972).

(See Peter Ladefoged, 'The interruptibility of speech'.)

1972 --, Stephen Krashen and Dale Terbeek. 'Is the left hemisphere specialized for speech, language or something else?' *Journal of the Acoustical Society of America* 51:79.

--, Stephen Krashen and Dale Terbeek. 'The left hemisphere is specialized for speech, language and something else' *Working Papers in Phonetics* 22:118-119.

1973 (See Peter Ladefoged, 'The interruptibility of speech'.)

1974 'The effects of psychological stress on speech: literature survey and background' *Journal of the Acoustical Society of America* 55:422.

--, Stepehn Krashen, Dale Terbeek, Roger Remington and Richard Harshman. 'Is the left hemisphere specialized for speech, language and/or something else?' *Journal of the Acoustical Society of America* 55:319-327.

--, and Peter Ladefoged. 'Two "voiceprint" cases' *Journal of the Acoustical Society of America* 55:463.

Roger Remington

1974 --, Stephen Krashen and Richard Harshman. 'A Possible sex difference in degree of lateralization of dichotic stimuli?' *Journal of the Acoustical Society of America* 55:434.

(See George Papcun, 'Is the left hemisphere specialized for speech, language and/or something else?')

Lloyd Rice

1971 'Glottal pulse waveform effects in line analog speech synthesis' *Working Papers in Phonetics* 19:48-54. Abstracted in *Journal of the Acoustical Society of America* 50:107.

'A new line analog speech synthesizer for the PDP-12' *Working Papers in Phonetics* 17:58-75.

'Notes on some recent computer programs' *Working Papers in Phonetics* 19:78-83.

1972 'Analysis of tongue parameters for vowels' *Working Papers in Phonetics* 23:32-36. Abstracted in *Journal of the Acoustical Society of America* 52:182.

1974 --, and Sven Ohman. 'On the relationship between formant bandwidths and vocal-tract shape features' *Journal of the Acoustical Society of America* 55:385.

Raymond Silverstein

1971 (See Peter Ladefoged, 'The interruptibility of speech'.)

1973 (See Peter Ladefoged, 'The interruptibility of speech'.)

Timothy S. Smith

1971 'A phonetic study of the function of the extrinsic tongue muscles' *Working Papers in Phonetics* 18.

Sarah Spitz

1973 (See Stephen Krashen, 'Language lateralization in a case of extreme psychosocial deprivation'.)

Robert P. Stockwell

1971 'The role of intonation: reconsiderations and other considerations' *Working Papers in Phonetics* 21:25-49.

Dale Terbeek

1971 --, and Richard Harshman. 'Cross language differences in the perception of natural vowel sounds' *Working Papers in Phonetics* 19:26-38. Abstracted in *Journal of the Acoustical Society of America* 50:147.

(See George Papcun, 'Is the left hemisphere specialized for speech, language or something else?')

1972 --, and Richard Harshman. 'Is vowel perception non-Euclidean?' *Working Papers in Phonetics* 22:13-29. Abstracted in *Journal of the Acoustical Society of America* 51:81.

(See George Papcun, 'Is the left hemisphere specialized for speech, language or something else?')

(See George Papcun, 'The left hemisphere is specialized for speech, language and something else'.)

Ralph Vanderslice

1971 --, and Peter Ladefoged. 'Binary suprasegmental features' *Working Papers in Phonetics* 17:6-24.

1972 --, and Peter Ladefoged. 'Binary suprasegmental features and transformational word-accentuation rules' *Language* 48:819-838.

Diana Van Lancker

1972 'Language lateralization and grammars' *Working Papers in Phonetics* 23:24-31. Abstracted in *Working Papers in Phonetics* 22:10-12.

'Language processing in the brain' *Working Papers in Phonetics* 23:22-23.

1973 --, and Victoria A. Fromkin. '"Tone" and pitch perception' *Journal of the Acoustical Society of America* 53:346.
Also published as 'Hemispheric specialization for pitch and "tone" evidence from Thai' *Journal of Phonetics* 1:101-109.

1975 'Heterogeneity in language and speech: neurolinguistic studies' *Working Papers in Phonetics* 29.

Theo Vennemann

1971 'The interpretation of phonological features in assimilation rules' *Working Papers in Phonetics* 19:62-68.

--, and Peter Ladefoged. 'Phonetic features and phonological features' *Working Papers in Phonetics* 21:13-24.
Also published in *Lingua* 32:61-74 (1973).

1973 --, and Peter Ladefoged. 'Phonetic features and phonological features' *Lingua* 32:61-74.