Reprinted from LANGUAGE
Vol. 46, No. 2, Part 1, June, 1970
Printed in U.S.A.

THE ARTICULATORY IMPLEMENTATION OF THE BREATH-GROUP AND PROMINENCE:

CRICO-THYROID MUSCULAR ACTIVITY IN INTONATION

PHILIP LIEBERMAN, MASAYUKI SAWASHIMA, KATHERINE S. HARRIS, AND THOMAS GAY

*Haskins Laboratories, New York City**

A theory has been proposed (Lieberman 1967) that accounts for some aspects of intonation in terms of two phonologic features, the BREATH-GROUP and PROMINENCE. Acoustic and physiologic correlates of these features were derived by experimental procedures that made use of subglottal air pressure and flow measurements as well as acoustic analysis. Perceptual data indicated that listeners 'decoded' certain intonational signals by means of 'motor theory perception' structured in terms of the 'archetypal', i.e., primary, articulatory correlates of these features. In the present study this theory was tested by recording the electrical activity of the crico-thyroid muscle of the larynx for a set of 480 short statements and yes-no questions that sometimes had non-terminal [+prominent] syllables. Independently derived data of Fromkin & Ohala 1968 also were examined, and were found to be consistent with the theory proposed by Lieberman except that [+prominent] syllables in UNMARKED BREATH-GROUPS had crico-thyroid activity. In yes-no questions where the crico-thyroid was active at the end of the MARKED BREATH-GROUP, non-terminal [+prominent] syllables had no crico-thyroid activity. The archetypal articulatory correlate of the marked breath-group is an increase in laryngeal tension; [+prominence] involves an increase in subglottal air pressure as well as increases in vowel duration and generally heightened muscular activity. Implementation rules relate the phonologic features to their archetypal and secondary articulatory correlates.

In a recent study (Lieberman 1967) some linguistic aspects of intonation were analysed in terms of two phonologic features, the BREATH-GROUP and PROMINENCE, which were defined in terms of their acoustic and physiologic correlates. Phonologic features may be regarded as psychological constructs that reflect the constraints of both the human speech-producing apparatus and auditory perception (Lieberman 1969). Some phonologic features are closely related to an articulatory maneuver that involves a specific muscle: for example, the feature NASALITY at the articulatory level is physically effected by means of the levator palatini muscle which closes or opens the nasal cavity to the rest of the vocal tract. However, many phonologic features cannot be related in an invariant manner to an articulatory maneuver that involves a particular muscle, muscle group, or anatomical structure. The phonologic feature STOP, for example, can be effected by means of any one of a number of articulatory maneuvers. The condition [+stop] at the articulatory level must be regarded as a 'state' function, i.e., a state of occlusion in the oral vocal tract.

Other features must also be regarded as state functions at the articulatory level. The prosodic features VOICING, PROMINENCE, and the BREATH-GROUP each involve the coördinated activity of many muscles; under different conditions, dif-

---

* Lieberman is also affiliated with the University of Connecticut, Storrs; Sawashima with the University of Tokyo; and Harris with the City University of New York.

ferent muscles may be involved in the implementation of a given feature. Indeed, the human larynx is so constructed that the fundamental frequency of phonation is a function of both the transglottal air pressure drop and the tensions of the laryngeal muscles (Muller 1848:1000–15, Van den Berg 1960, Ladefoged 1962, Flanagan & Landgraf 1968, Lieberman 1967, Lieberman et al. 1969). In the absence of any information save that contained in the acoustic signal, it would seem difficult to ascertain whether a change in fundamental frequency is due to a change in the tension of the laryngeal muscles or to a change in the subglottal air pressure. That is to say, there are two different mechanisms, laryngeal and subglottal respiratory maneuvers, that can produce similar changes in fundamental frequency. However, under certain conditions, listeners appear to 'decode' intonational signals in terms of the articulatory mechanisms that could underlie the acoustic signal (Lieberman 1967). The listeners, in other words, appear to be acting as though they make use of a 'motor theory' of speech perception to interpret the acoustic signal.[1] It is important to note at this point that a motor theory of speech perception does not mean that listeners consciously or unconsciously mimic the sounds that they hear, nor does it even imply that the listeners actually 'compute' the acoustic effects of various articulatory maneuvers, using a technique of 'analysis-by-synthesis' (Halle & Stevens 1959).

A motor theory of speech perception, in the sense that we will use the term, simply states that the perceptual recognition routines that are used when speech is decoded are structured in terms of the constraints of the human vocal apparatus. Thus, just as a particular species of frog has auditory receptors that respond to the spectrum of that species' mating croak (Capranica 1965), humans apparently have a perceptual mechanism that responds to the formant transitions of the stops /b d g/ (Liberman et al. 1967).

We still have not answered the apparent paradox that seems to apply to a motor theory mode of perception for intonation. If alternate articulatory mechanisms can produce the same fundamental frequency variations, how can listeners decode the intonational signal in terms of the underlying articulatory maneuvers? The solution proposed in Lieberman 1967 was that the listeners decode in terms of the 'archetypal' pattern of articulatory activity. The archetypal pattern was defined as the simplest, or basic, state of muscular control that would produce the intonational signal. The hypothetical archetypal NORMAL BREATH-GROUP was said to involve a state of minimal laryngeal control throughout expiration so that changes in fundamental frequency (excluding those due to mechanical and aerodynamic interactions with the rest of the vocal tract) would follow from changes in transglottal air pressure. The fundamental frequency of phonation would thus fall rapidly at the end of a breath-group where the subglottal air pressure must change from a positive to a negative pressure in order to get air into the lungs during inspiration. The normal breath-group thus would be the basis of the intonation contour transcribed by Jones 1932 and by Armstrong & Ward 1926 as Tune I, by Pike 1945 as the 'pause' [//], and by Trager & Smith 1951 as 231 #.

---

[1] Lehiste & Peterson 1959 propose that '... the perception of linguistic stress is based upon judgments of the physiological effort involved in producing vowels'. Ladefoged 1962 relates the perception of stress to judgments of subglottal air pressure.

The data presented by Lieberman 1967 showed that some adult speakers used the archetypal pattern of articulatory activity to produce these intonational signals. It was not possible at the time to measure the electrical activity of the laryngeal muscles. Laryngeal maneuvers thus had to be inferred from measures of subglottal air pressure, lung volume, and the acoustic signal.

These data further indicated that the MARKED BREATH-GROUP—[+breath-group]—involved a change in laryngeal tension at the end of the breath-group which offset the falling subglottal air pressure, producing either a rising or level terminal fundamental frequency contour. The marked breath-group appeared to be the basis for the signals that have been transcribed by Jones and by Armstrong & Ward as Tune II, by Pike as the 'pause' [/], and by Trager & Smith as either 232| or 232||.[2] The acoustic correlates of the phonologic feature PROMINENCE included duration, fundamental frequency, and intensity (Fry 1955, Lieberman 1960). The identifiable articulatory correlates included increasing the duration of a segment and increasing the subglottal air pressure (Ladefoged 1962).

The analysis of intonation presented in Lieberman 1967 also was consistent with the notion of motor theory decoding (Liberman et al. 1967) of certain intonation contours, in particular of [+breath-groups] that had a non-terminal [+prominent] syllable. Note that the two intonation contours schematized in Figure 1 have different terminal rises. The listeners, in a carefully controlled psycho-acoustic experiment (Hadding-Koch & Studdert-Kennedy 1964), said that both of these contours had the same terminal rise. The listeners obviously were not merely responding to the physically present fundamental frequency signal; they instead seemed to be evaluating the terminal fundamental frequency contours in terms of the degree of laryngeal tension that would be present in natural speech, and to be interpreting the non-terminal fundamental frequency peaks as though they were a consequence of a peak in the subglottal air pressure function (Ladefoged 1962). The listeners also appeared to 'know' that a non-terminal peak in the subglottal air pressure function will result in a lower subglottal air pressure at the end of the breath-group. This 'air pressure perturbation' hypothesis is, of course, part of this theory for the production and perception of intonation. Note that the intonation contour in Figure 1 that had the greater non-terminal peak fundamental frequency has the smaller terminal fundamental frequency. Since the lower contour had a higher non-terminal peak f_0, a greater non-terminal peak subglottal air pressure would have been employed than for the upper contour. This would result in a lower subglottal air pressure at the end of the breath-group for the lower contour, and an equivalent degree of laryngeal tension would thus result in a lower terminal fundamental frequency.

The listener's perceptual recognition routine for intonation, in other words, 'knows' that:

(a) terminal non-falling fundamental frequency contours are the result of increases in laryngeal tension;

---

[2] The phonetic distinction between the terminals [|] and [||] may be a co-articulation effect. When a speaker uses a [+breath-group] in a non-sentence-final position, he may not complete the tensioning of his laryngeal muscles at the end of the [+breath-group] because he must rapidly begin to relax these muscles for the beginning of the breath-group that follows. Similar effects occur for other phonologic features; see Lindblom 1963 for a careful study of co-articulation effects in vowels.

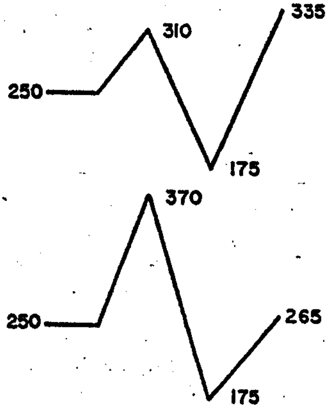

FIGURE 1. Two fundamental frequency contours that listeners identified as having the 'same' terminal fundamental frequency rise (from Hadding-Koch & Studdert-Kennedy 1964).

(b) non-terminal fundamental frequency peaks—at least in marked breath-groups—are the result of a momentary increase in subglottal air pressure;

(c) the presence of a non-terminal peak in subglottal air pressure will result in a lower terminal subglottal air pressure than would be the case if the non-terminal peak were absent;

(d) all these maneuvers modify an archetypal normal breath-group.

The theory that we have briefly summarized is fairly complex, though it is no more complex than the theory that we must propose for the perception of the stop consonants (Liberman et al. 1967). We have said that phonologic features may have complex articulatory correlates, and we have further stated that complex perceptual recognition routines are involved in decoding the acoustic signal in terms of the articulatory correlates of these features.

Recent papers by Ohala & Hirano (1967), by Vanderslice (1967), and by Fromkin & Ohala (1968), and a review by Kim (1968) dispute both the general nature of the phonologic features that we have proposed and the specific correlates of the archetypal breath-group and prominence that we have discussed, as well as the status of perception structured in terms of the constraints of speech production, i.e. the motor theory of speech perception. We will present some new electromyographic data obtained from laryngeal and supralaryngeal muscles during speech. These data are relevant both to the general question of the relationship between a phonologic feature and a particular muscular maneuver and to the specific question of the status of the analysis of intonation proposed by Lieberman 1967. We will also discuss the data presented by Fromkin & Ohala, which are consistent with, and which complement, our new data.

EXPERIMENTAL DATA. Electromyographic data were obtained from the crico-thyroid and orbicularis muscles of a female speaker of American-English (KSH). Concentric needle electrodes were inserted into these muscles, and the electrical muscle potentials were amplified and recorded on a multi-channel magnetic tape recorder while they were being monitored on a Grass Instruments oscillograph. The speaker throughout this experiment spoke short sentences like the following (where the small capitals indicate emphasis):

It is pattering.
Is it pattering?
It is PATTERING.
Is it PATTERING?

The words *pattering*, *packer*, *parfait*, *keeper*, and *recap* were produced in unemphasized declarative sentences, in unemphasized yes-no questions, and in declarative and yes-no questions with emphasis. The speaker read a list in which these sentences appeared in a random order until each sentence was uttered twenty times; 480 utterances were thus recorded. A computer program was then used which averaged the electromyographic signals for each group of twenty tokens of each sentence, excluding those few tokens of each utterance which were faulty in various ways. The computer program 'lined up' each sentence in terms of the cessation of phonation. The end of phonation for each sentence was determined from the acoustic signal which was simultaneously recorded on both the magnetic tape and the oscillogram.

In Figure 2a we have presented the integrated and averaged electromyographic

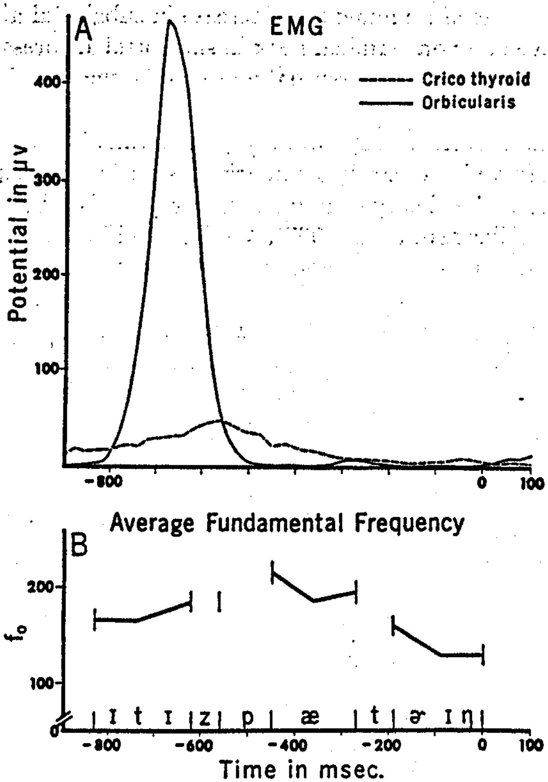

FIGURE 2. Averaged electrical activity of the orbicularis and crico-thyroid muscles (a) and fundamental frequency (b) for the sentence *It is pattering*. Note that the crico-thyroid muscle, which effects fundamental frequency changes, is not active at the end of the sentence ($t = 0$).

data for the sentence *It is pattering*. The integral of the action potentials of a muscle is a valid measure of muscular activity. The averaging technique, which averages a number of examples of the same linguistic construct, lessens the possibility of artifact influence on our conclusions. In this case, 20 tokens of the sentence were averaged together. The activity of the orbicularis muscle is plotted in a solid line, while the activity of the crico-thyroid muscle is plotted in dashes. The vertical, ordinate scale on the graph is the electrical activity of the muscles in microvolts. Time is plotted on the horizontal axis in milliseconds. The negative and positive values of time refer to the computer 'line up' point.

In Figure 2b, the fundamental frequency of phonation has been averaged for six of the above utterances and plotted on the same time scale. The fundamental frequency was measured on narrow band spectrograms by tracking the fifth harmonic of the fundamental.

Note that the electrical activity of the orbicularis muscle is limited to a single peak centered on $t = -650$ msec. This peak corresponds to the closure of the lips for /p/ (Harris et al. 1965, 1968), as we can see by referring to the fundamental frequency contour in Figure 2b, which bears a phonetic transcription of the utterance. Note that no electrical activity occurs in the crico-thyroid muscle at the end of the sentence ($t = 0$ msec.). A small (50 microvolt) peak in crico-thyroid activity does occur at $t = -550$ msec., and corresponds with the primary lexical stress of the word *pattering*.

In Figure 3a similar data are plotted for the sentence *Is it pattering?* Note that a peak in the orbicularis channel still occurs,[3] and that, in contrast to Figure 2a, electrical activity occurs in the crico-thyroid muscle at the end of the sentence. The activity of the crico-thyroid muscle results in an increase in the fundamental frequency of phonation (Muller 1848, Van den Berg 1960, Ohala & Hirano 1967, Öhman 1967).

In Figure 3b the fundamental frequency of phonation has been averaged for six of the above utterances and plotted on the same time scale. Note that the fundamental frequency also rises at the end of the breath-group. Examination of the data for all of the sentences showed that the increase in the electrical activity of the crico-thyroid muscle and the increase in the fundamental frequency occurred at the end of the marked breath-group, regardless of the number of syllables in the final word of the sentence or their lexical stress.

Averaged electromyographic data are presented in Figure 4 for five marked breath-groups. Note that the activity of the crico-thyroid is similar for all the questions. The change in stress pattern and the number of syllables have no effect on the observed crico-thyroid activity.

In Figure 5a, electromyographic data is presented for the word *pattering*, with overstress, in the sentence *It is* PATTERING. Note the presence of crico-thyroid activity centered at -600 msec., and the absence of crico-thyroid activity at the end of the breath-group. In Figure 5b, electromyographic data are presented for *pattering* as it occurred when it was overstressed in a MARKED breath-group, i.e.,

---

[3] The difference in size of the orbicularis peak between statement and question, on the basis of a more detailed analysis of statement-question contrasts, appears to be a random fluctuation.

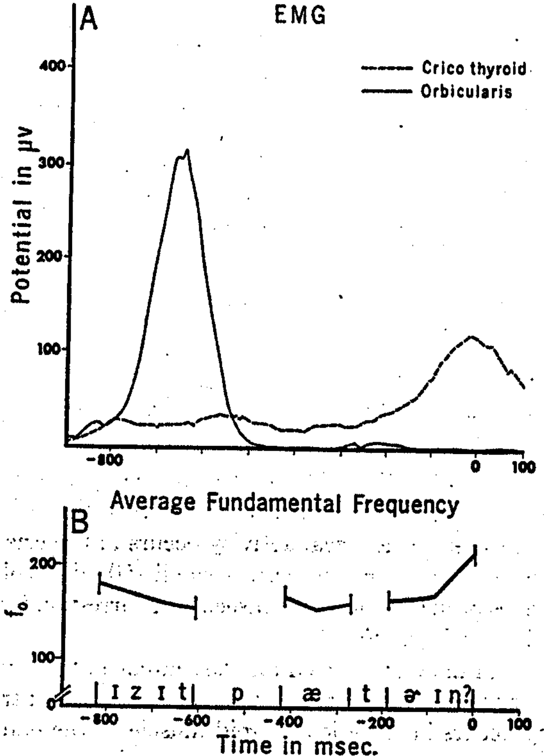

FIGURE 3. Averaged electrical activity of the orbicularis and crico-thyroid muscles (a) and fundamental frequency (b) for the sentence *Is it pattering?* Note that the crico-thyroid muscle is active at the end of the sentence ($t = 0$).

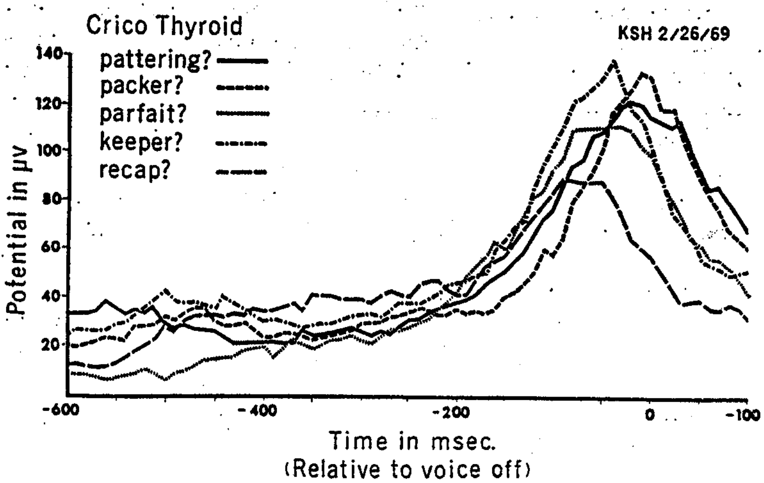

FIGURE 4. Crico-thyroid activity for five sentences (approximately 20 tokens per average) that each had the form *Is it pattering?*, *Is it a packer?* etc. All of the sentences have been 'lined up' with respect to the end of phonation ($t = 0$).

in the sentence *Is it* PATTERING? Note the presence of crico-thyroid activity at the end of the breath-group and the relative absence of crico-thyroid activity at −600 msec. Our acoustic analysis of the sentences plotted in Figure 5 revealed that the emphasis placed on the word *pattering* was manifested by the acoustic correlates of the phonetic feature prominence that have been described by Jones 1932, Fry 1955, Bolinger 1958, Lieberman 1960, and others. The duration of the vowels of *pattering* was sometimes greater when it was marked by prominence. We did not measure the intensity of the speech signal, but in all likelihood it was probably somewhat greater for the [+prominent] examples.

DISCUSSION. Let us summarize the observations drawn from this experimental data. About 480 short sentences that were uttered by a single speaker were analysed. Some of these sentences were yes-no questions produced by means of a MARKED BREATH-GROUP. Some of the sentences had words that the speaker was asked to emphasize. The sentences with emphasized words had non-terminal [+prominent] syllables. The data showed the following:

(1) The UNMARKED BREATH-GROUPS terminated with a falling f_0 contour, and there was no crico-thyroid activity at the end of these breath-groups.

(2) The MARKED BREATH-GROUPS were always terminated by a rising f_0 contour produced by means of increased tension of the crico-thyroid muscle.

(3) When a syllable was marked by [+prominence] in an unmarked breath-group, it had a higher f_0, and there was increased activity of the crico-thyroid muscle that was correlated with the particular syllable.

(4) When a syllable was marked with [+prominence] in a marked breath-group, it also had a higher f_0. However, NO increase in crico-thyroid muscle activity could be correlated with the f_0 prominence of the syllable.

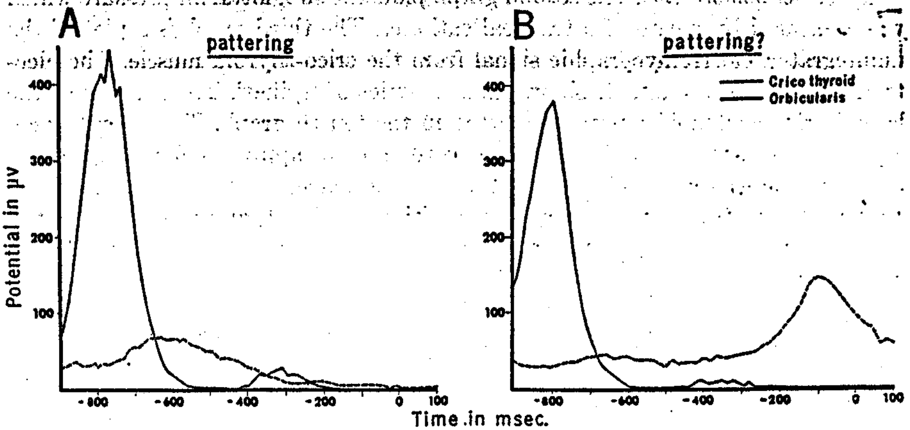

FIGURE 5. Averaged electrical activity of the orbicularis and crico-thyroid muscles for the sentences (a) *It is* PATTERING and (b) *Is it* PATTERING? Emphasis was placed on the word *pattering* in both the declarative sentence and the yes-no question. Note that a peak in crico-thyroid activity is evident centered at $t = -600$ msec in (a), whereas no non-terminal peak occurs in (b) where the crico-thyroid is active at the end of the marked breath-group.

(5) Some crico-thyroid activity could be correlated with primary lexical stress in an unmarked breath-group. However, the magnitude of this crico-thyroid activity was 50 percent smaller than was the case for the [+prominent] syllables that occurred in unmarked breath-groups.

Our interpretation of this data is that the primary, i.e., archetypal articulatory correlate of the [+breath-group] is an increase in crico-thyroid muscular activity (and possibly other synergetic laryngeal muscular activity).[4] In contrast, the archetypal correlate of [+prominence] is an increase in subglottal air pressure.[5] In an unmarked breath-group, where the speaker is not going to do anything with his crico-thyroid muscle at the end of the breath-group, he may use it to manifest [+prominence] together with increased subglottal air pressure. In a [+breath-group] the crico-thyroid muscle is 'reserved' to implement the terminal f_0 contour; the speaker thus uses an increase in subglottal air pressure and/or an increase in vowel duration to implement the non-terminal [+prominence]. The speaker may perhaps use less air pressure for a [+prominence] that occurs in a [−breath-group] where he can use laryngeal maneuvers to implement the [+prominence] in addition to the air pressure peak. These data are, of course, consistent with the analysis-by-synthesis motor theory model proposed in Lieberman 1967. The listeners in the Hadding-Koch & Studdert-Kennedy psychoacoustic experiment appeared to decode the intonational signals as though prominence were caused by a subglottal air pressure peak. The data are, moreover, consistent with the electromyographic data that have been published by Fromkin & Ohala.

We have reproduced some of the Fromkin & Ohala data in Figures 6 and 7. We have taken the liberty of marking the columns of these figures since it will make the discussion easier to follow. First, note the data in column A of Figure 6. The first graph of this column is a plot of the fundamental frequency of the utterance, *Bev bombed Bob*. The second graph plots the subglottal air pressure which was measured by means of a tracheal catheter. The third graph is a plot of the unintegrated electromyographic signal from the crico-thyroid muscle. The electrical activity of the muscle shows up as a series of 'spikes'. The activity of the lateral crico-arytenoid muscle is plotted in the fourth graph. The lateral crico-arytenoid is used to adduct the vocal cords and to apply medial compression (Van den Berg). In order for phonation to take place, the vocal cords must be adducted from their open, respiratory position. The fifth graph is a plot of the acoustic signal, while the bottom graph is a timing pulse that indicates an interval of 100 msec.

---

[4] There is some evidence to indicate that the laryngeal muscles must act in concert to effect changes in fundamental frequency (Van den Berg; Ohala & Hirano).

[5] Ladefoged 1968 notes that subglottal respiratory activity is always correlated with [+prominence] in his data whereas other articulatory maneuvers are not. Ladefoged (personal communication) would make subglottal activity a necessary articulatory correlate of [+prominence]. This is, of course, a stronger claim than stating that subglottal activity is the archetypal articulatory correlate of [+prominence]. However, in either case laryngeal maneuvers are secondary correlates of [+prominence]. The secondary articulatory correlates (which include activity throughout the vocal tract) all appear to involve increased muscular activity (Harris et al. 1968; Lieberman 1969).

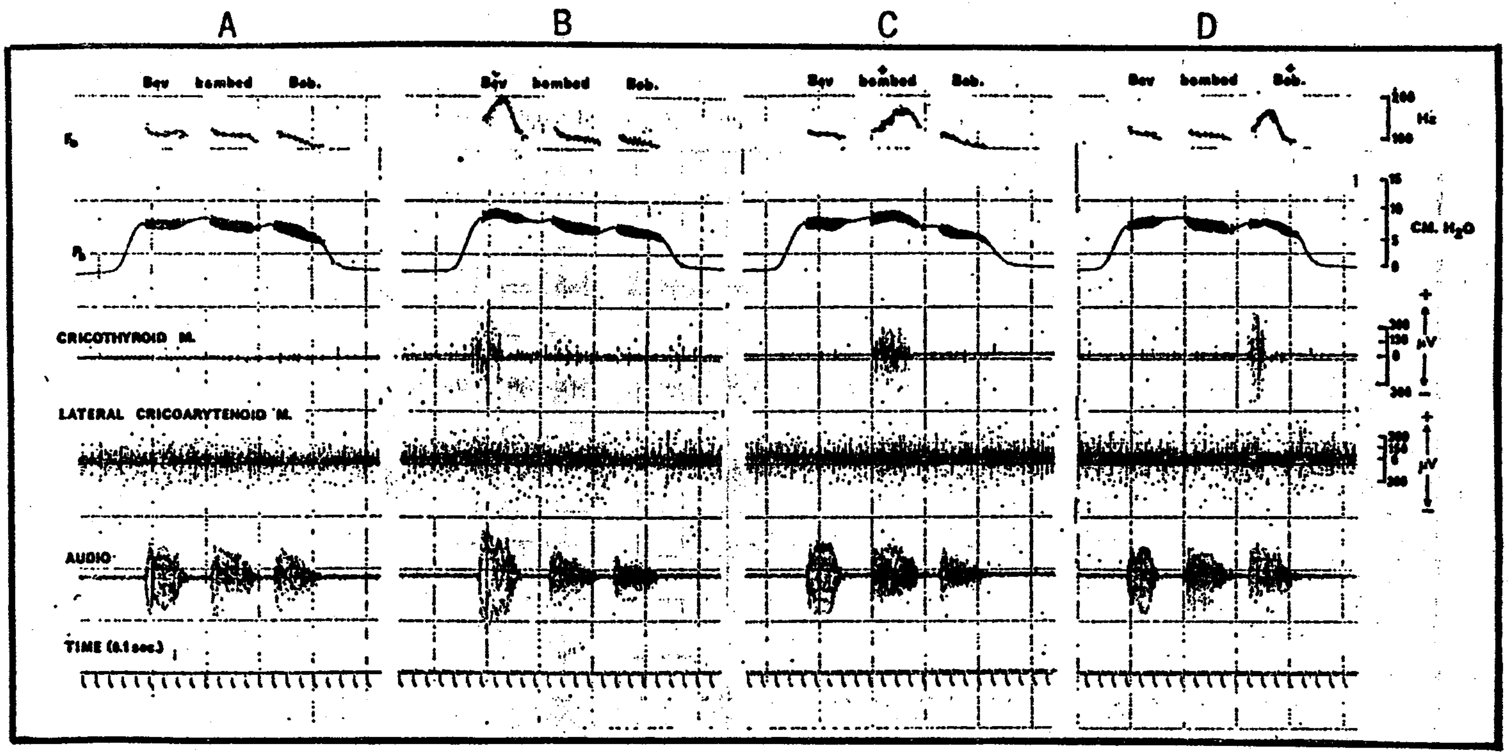

FIGURE 6. Electromyographic, physiologic, and acoustic data from Fromkin & Ohala 1968.

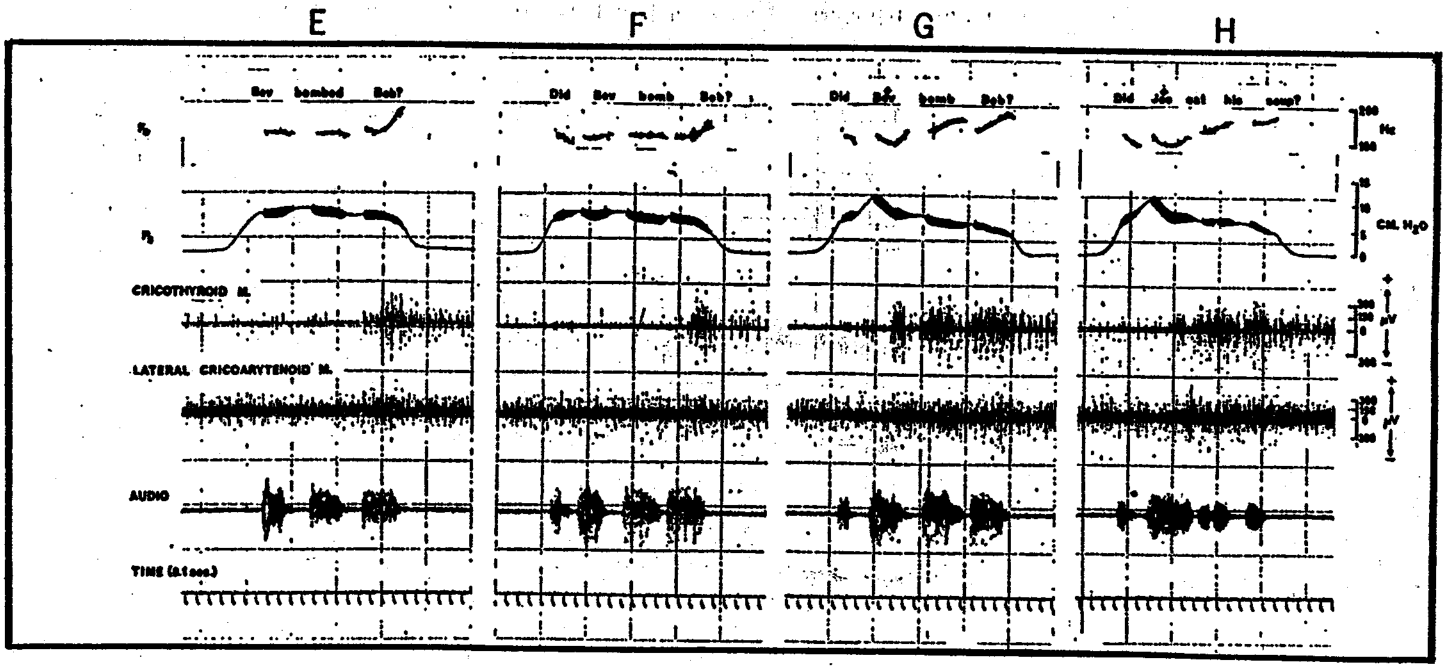

FIGURE 7. Electromyographic, physiologic, and acoustic data from Fromkin & Ohala 1968.

Note that the lateral crico-arytenoid muscle (the fourth graph in Figure 6a) is firing both before and after the subglottal air pressure builds up and decays. The subglottal air pressure curve thus would appear to reflect subglottal articulatory maneuvers. The entire utterance is voiced, so we should expect no perturbations in the f_0 function that were due to opening and closing the glottis for voiced-voiceless distinctions. Note that if the f_0 contour is traced and superimposed on the P_s contour, they are remarkably similar. There are falls in f_0 before voiced obstruents, but these falls are due to the supralaryngeal obstruction which, of course, reduces the transglottal air flow (Öhman). This is indeed a better example of a speaker who employs the archetypal articulatory correlates of the unmarked breath-group than any example in Lieberman 1967. The clever use of voiced phonetic material by Fromkin & Ohala removed perturbations of f_0 that may result from the 'setting and resetting' of laryngeal tension that must occur when voiced-voiceless distinctions occur (Lieberman 1968). The particular phonetic material which does not involve vowels with low first formants also removes one source of interaction between the supralaryngeal vocal tract and the larynx (Flanagan & Landgraf).[6] The changes in air pressure account for virtually all of the controlled changes in fundamental frequency in Figures 6 and 7, except for those that are the consequence of crico-thyroid activity. Fundamental frequency otherwise follows air pressure for all of the *Bev bombed Bob* contours except for parts of contours 6D, 7F, and 7G where it is evident that some other laryngeal muscular activity affects the f_0 contour. In all other instances, the changes in f_0 appear to be the result of the transglottal air pressure drop changing the fundamental frequency at an average rate of approximately 12.5 Hz/cm H_2O. The actual rate of change of f_0 with respect to air pressure varied between 6 and 20 Hz/cm H_2O in this data sample. This range of variation is consistent with Van den Berg, Flanagan & Landgraf, and Lieberman et al.[7]

---

[6] The rate of change of fundamental frequency with respect to changes in transglottal air pressure is a function of both the 'mode' of phonation and the configuration of the supralaryngeal vocal tract. Maximum sensitivity to air pressure occurs for vowels with low first formant frequencies like [u] and [i]. The rate of change of fundamental frequency with respect to air pressure varies from about 2.5 Hz/cm H_2O to 20 Hz/cm H_2O. The average fundamental frequency is thus a function of the mode of phonation (the manner in which the vocal cords collide, the balance of tissue and aerodynamic forces, etc.) and the vocal tract configuration (Peterson & Barney 1952; Flanagan & Landgraf; Lieberman et al.)

[7] Vanderslice, Kim, and Fromkin & Ohala claim that virtually all controlled changes in fundamental frequency are the result of laryngeal muscular activity. The effects of air pressure on fundamental frequency are supposed to be very slight (less than 3 to 5 Hz/cm H_2O). It is difficult to see how the rather vehement claims made by these authors can be reconciled with the data that they cite. It is evident that the average rate of change of fundamental frequency with respect to air pressure is about 12.5 Hz/cm H_2O in the data of Fromkin & Ohala, which form the main experimental basis of these claims. The fundamental frequency contour in Figure 6A is, for example, completely specified by the air pressure function. It is actually a better example of an adult speaker employing the archetypal, i.e. primary articulatory correlates of the [−breath-group] than any example in Lieberman 1967. This is probably due to the fact that the utterance *Bev bombed Bob* is to a normal sentence as a nonsense syllable like /bub/ recorded in an experiment is to a real word uttered in discourse. Whereas a speaker in normal discourse will deviate from articulatory gestures that he 'knows' he should use, he will be quite careful when he records 'test' syl-

In utterances 6B, C, and D, [+prominence] is manifested by means of both subglottal air pressure and crico-thyroid activity. In utterances 7E and F, the terminal f_0 contours of the marked breath-groups are also due to crico-thyroid activity, as is the case for utterances G and H.[8] However, note that the increase in crico-thyroid activity in G and H occurs about 200 msec. after the peak in P_s which is coincident with the marked f_0 prominences. The f_0 peaks associated with [+prominence] in these utterances apparently are a consequence of the air pressure peaks. Fromkin & Ohala's data is thus consistent with the theory proposed in Lieberman 1967 for [+prominence] in a [+breath-group]. Note too that these air pressure peaks are substantially greater than those that occur in utterances B, C, and D. Note that the 'air pressure perturbation' effect discussed by Lieberman 1967 is evidenced by these data. Compare the terminal air pressures of utterances G and H with F and E. The air pressure is lower (by about 2.5 cm, though it's hard to read the compressed scale accurately) at the end of utterances G and H. This follows from the probable programming of the respiratory system for a breath-group (cf. Lieberman 1967:54, 71, 98–100).

PHONOLOGIC FEATURES AND ARTICULATORY MANEUVERS. The data that we have discussed demonstrate clearly that phonologic features and articulatory maneuvers cannot be mapped in an invariant one-to-one manner. The feature [+prominence] involved activity of the crico-thyroid muscle only when the speakers 'knew' that they would not be using this muscle at the end of a [+breath-group]. The implementation rules of the phonetic component that relate the phonologic feature [+prominence] to the articulatory output thus must note whether a segment marked [+prominent] occurs in a [+breath-group]. The implementation rule, in other words, will assign different muscular maneuvers to the state [+prominent] in different contexts. Of course there is nothing particularly novel in this concept. The phonologic feature [+stop] will have very different articulatory correlates for labial and glottal stops. Phonologic features,

---

lables in an experiment. When a speaker produces a 'nonsense' sentence like *Bev bombed Bob*, he is quite careful and thus tends to use the archetypal or primary articulatory correlates of the [−breath-group]. Fromkin & Ohala's data thus, in general, replicate Lieberman 1967.

[8] Note that the crico-thyroid activity and terminal fundamental frequency rise occur earlier in these utterances than is the case for the utterances plotted in Figures 3, 4, and 5b. This may reflect speaker-specific differences in the implementation rules that relate phonologic features to muscle commands (Lieberman 1969); that is, different speakers may have slightly different timing for the onset of crico-thyroid activity relative to the end of the [+breath-group]. The earlier onset of crico-thyroid activity could, however, reflect the presence of a different phonologic feature. Bolinger discusses intonation contours that may call for the introduction of features that supplement the breath-group and prominence. More data is necessary to resolve this question. Note also that the speaker in Figures 7F, G, and H started each utterance with a falling f_0 contour though subglottal air pressure was rising. Similar effects were noted in Lieberman 1967. These effects must, of course, be due to laryngeal muscular activity (perhaps the vocalis muscle). They thus must reflect either various implementations of the breath-group that differ from its archetypal muscular correlates, or additional phonologic features. Systematic experiments which involve both the analysis of speech and psycho-acoustic tests with synthetic speech are in progress to explore these possibilities, as well as the possible effects of extra-linguistic factors like emotion.

in general, appear to be 'state' functions insofar as the physical invariance in each feature is a 'state' of the articulatory apparatus, e.g. an oral vocal tract occlusion for [+stop], or the state of phonation or incipient phonation for [+voicing] (Lieberman 1969). It is only in certain cases like [+nasal] that an invariant articulatory maneuver can be assigned to a particular feature.

The implementation rules that assign laryngeal maneuvers to [+prominence] are, however, of particular interest since they must take into account articulatory events that may take place several syllables after the [+prominent] syllable. Lashley 1951 stated that associative theories cannot account for processes like the production of phonetic sequences of speech. Wickelgren 1969 proposed that speech could be encoded by means of associations between context-sensitive elementary motor responses. The word 'stop' would thus be encoded 'allophonically' as $/{}_{\#}s_{t}\ {}_{s}t_{o}\ {}_{t}o_{p}\ {}_{o}p_{\#}/$ rather than 'phonemically' as /s t o p/. Wickelgren's theory of course requires a greater 'memory' in which context-sensitive allophones would be stored. The memory requirements are however plausible in terms of human mental capacity so long as only the immediate context of each allophone must be considered. Effects like the manifestations of crico-thyroid muscular activity in [+prominent] syllables which involve the 'remote' context of the [+prominent] syllable argue against Wickelgren's context-sensitive allophonic encoding of speech. The memory requirements would preclude simple associative encoding.

ARCHETYPAL ARTICULATORY CORRELATES. The data that we have been discussing suggest how the archetypal or primary articulatory correlates of phonologic features may be manifested. The articulatory bases of the phonologic feature PROMINENCE must obviously include laryngeal maneuvers. The archetypal articulatory correlate of *prominence*, however, appears to be a peak in subglottal air pressure (Ladefoged 1962, 1968; Lieberman 1967). The data support, in detail, the analysis proposed in Lieberman 1967 if prominence is regarded as a state function (Lieberman 1969). The data also support the motor theory decoding of intonational signals that is proposed in Lieberman 1967.

OTHER PROSODIC FEATURES. The analysis proposed in Lieberman 1967 is limited to only two prosodic features, the breath-group and prominence. It is apparent that other prosodic features also must be considered in order to account for phenomena like the accent system of Swedish (Öhman) or the tone systems that occur in many languages (Wang 1967). Still other features may be necessary to account for the abrupt fundamental frequency falls that serve as phonetic manifestations of contrast and emphasis in English (Bolinger). Some of the data that we have presented—for example, the crico-thyroid activity associated with lexical stress in the unmarked breath-groups—may be articulatory manifestations of features like those proposed by Bolinger, Öhman, and Wang. The data of Fromkin & Ohala also show muscular activity that may reflect the presence of features other than the breath-group and prominence. There also may be complex interactions between these unspecified features and the breath-group and prominence. Further experiments and theoretical refinements are obviously necessary. However, the data that we have discussed are consistent with the framework for the analysis of intonation that we have proposed and indicate that we may be going in the right direction.

REFERENCES

ARMSTRONG, LILIAS E., and IDA C. WARD. 1926. Handbook of English intonation. Leipzig and Berlin: Teubner.

BOLINGER, DWIGHT. 1958. A theory of pitch accent in English. Word 14.109–49.

CAPRANICA, ROBERT R. 1965. The evoked vocal response of the bullfrog. Cambridge, Mass.: MIT Press.

FLANAGAN, JAMES L., and LOIS LANDGRAF. 1968. Self-oscillating source for vocal-tract synthesizers. IEEE Trans. Audio Electroacoustics AU-16.57–64.

FROMKIN, VICTORIA, and JOHN OHALA. 1968. Laryngeal control and a model of speech production. Working papers in phonetics, UCLA, 10.98–110.

FRY, DENNIS B. 1955. Duration and intensity as physical correlates of linguistic stress. J. Acoust. Soc. Am. 35.765–9.

HADDING-KOCH, KERSTIN, and MICHAEL STUDDERT-KENNEDY. 1964. An experimental study of some intonation contours. Phonetics 11.175–85.

HALLE, MORRIS, and KENNETH N. STEVENS. 1959. Analysis by synthesis. Proc. of Seminar on Speech Compression and Processing, Air Force Cambridge Research Center, AFCRC-TR-59-198.

HARRIS, KATHERINE S.; GLORIA F. LYSAUGHT; and MALCOLM M. SCHVEY. 1965. Some aspects of the production of oral and nasal labial stops. Language and Speech 8.135–47.

HARRIS, KATHERINE S.; THOMAS GAY; GEORGE N. SHOLES; and PHILIP LIEBERMAN. 1968. Some stress effects on electromyographic measures of consonant articulation. Haskins Laboratories status report, SR 13/14.137–52, New York City. (Also presented at the Kyoto Speech Symposium of the 6th Intl. Cong. on Acoustics, 29 August 1968.)

JONES, DANIEL. 1932. An outline of English phonetics, 3rd ed. New York: Dutton.

KIM, CHIN-WU. 1968. Review of Intonation, perception and language, by P. Lieberman. Lg. 44.830–42.

LADEFOGED, PETER. 1962. Subglottal activity during speech. Proc. Fourth Intl. Cong. Phonetic Sciences, 247–65.

——. 1968. Linguistic aspects of respiratory phenomena. Annals of the New York Academy of Sciences, 155.141–51.

LASHLEY, KARL S. 1951. The problem of serial order in behavior. Psycholinguistics, ed. by S. Saporta, 180–98. New York: Holt, Rinehart & Winston.

LEHISTE, ILSE, and GORDON E. PETERSON. 1959. Vowel amplitude and phonemic stress in American English. J. Acoust. Soc. Am. 31.428–35.

LIBERMAN, ALVIN M.; FRANKLIN S. COOPER; DONALD P. SHANKWEILER; and MICHAEL STUDDERT-KENNEDY. 1967. Perception of the speech code. Psychological Review 74.431–61.

LIEBERMAN, PHILIP. 1960. Some acoustic correlates of word stress in American English. J. Acoust. Soc. Am. 32.451–4.

——. 1967. Intonation, perception and language. Cambridge, Mass.: MIT Press.

——. 1968. Vocal cord motion in man. Annals of the New York Academy of Sciences, 155.28–38.

——. 1969. On the physical bases of phonologic features. To appear in Langages. (Abstract in Handbook of the 43rd Annual Meeting of the LSA, p. 62, New York City.)

LIEBERMAN, PHILIP; RONALD KNUDSON; and JERE MEAD. 1969. Determination of the rate of change of fundamental frequency with respect to subglottal air pressure during sustained phonation. J. Acoust. Soc. Am. 45.1537–43.

LINDBLOM, BJORN. 1963. On vowel reduction. Speech Transmission Laboratory Rept. 29, Royal Institute of Technology, Stockholm, Sweden.

Muller, Johannes. 1848. The physiology of the senses, voice and muscular motion with the mental faculties, trans. by W. Baly. London: Walton & Maberly.

Ohala, John, and Minoru Hirano. 1967. Studies of pitch change in speech. Working papers in phonetics, UCLA, 7.80–4.

Öhman, Sven E. G. 1967. Word and sentence intonation: a quantitative model. Speech Transmission Laboratory, QPSR 2–3/1967.20–55, Royal Institute of Technology, Stockholm, Sweden.

Peterson, Gordon E., and Harold L. Barney. 1952. Control methods used in a study of the vowels. J. Acoust. Soc. Am. 24.175–84.

Pike, Kenneth L. 1945. The intonation of American English. Ann Arbor: Univ. of Michigan Press.

Trager, George L., and Henry L. Smith. 1951. Outline of English structure. (Studies in Linguistics, Occasional papers, 3.) Norman, Oklahoma: Battenburg.

Van den Berg, Jan. 1960. Vocal ligaments versus registers. Current Problems in Phoniatrics and Logopedics, 1.19–34.

Vanderslice, Ralph. 1967. Larynx versus lungs: Cricothyrometer data refuting some recent claims concerning intonation and archetypality. Working papers in phonetics, UCLA, 7.69–80.

Wang, William S.-Y. 1967. Phonological features of tone. IJAL 33.93–105.

Wickelgren, Wayne A. 1969. Context-sensitive coding, associative memory and serial order in (speech) behavior. Psychological Review 76.1–15.

[Received 16 June 1969]

Since this article was originally written we have obtained additional data using variations of the same electromyographic procedures on additional speakers. These data together with extensive analysis and discussion, both among ourselves and with colleagues, indicate that our original analysis oversimplifies the articulatory implementation of [+prominence]. Prominence appears to be implemented partly by varying laryngeal tension, partly by varying duration, and partly by varying subglottal air pressure. Another way of saying this is that the implementation of *+prominence* is variably mapped out onto physiological states by different speakers. The description of the Fromkin & Ohala data show quite clearly that the fundamental frequency contour is affected by both subglottal and glottal activity. However, as we originally noted, there are some changes in the fundamental frequency contour that cannot be explained either as responses to the action of the muscles being monitored or the subglottal air pressure.