Offprint from *CONSONANT TYPES AND TONE*
Southern California Occasional Papers
in Linguistics No. 1, July 1973
Edited by Larry M. Hyman

THE PHYSIOLOGY OF TONE

BY

JOHN J. OHALA

DEPARTMENT OF LINGUISTICS
UNIVERSITY OF CALIFORNIA
BERKELEY, CALIFORNIA 94720

Published by the Linguistics Program
University of Southern California
Los Angeles, California 90007

0. INTRODUCTION

The sound patterns of tone, as described in the phonological literature--much of it very recent (Bien-Ming 1931, Haudricourt 1954, 1961, Wang 1967, Greenberg 1970, Chang 1972, Hyman and Schuh 1972, Cheng 1973), exhibit many interesting assymmetries which are determined by phonetic context, for example:

1. Historically, many consonants seem to have induced tonal differences on the vowel following, but few consonants, primarily only [h] and [ʔ], seem to have caused tonal changes on the vowel preceding.*
2. Specifically, [ɦ], voiced or breathy-voiced obstruents seem to depress tone on the vowel following, whereas voiceless, particularly voiceless aspirated obstruents, [h], and implosives seem to raise tone on the following vowel. There is also some evidence that [ʔ] and [h] may raise and lower, respectively, the tone of the preceding vowel (cf. Matisoff 1973).*
3. Falling tones outnumber rising tones.
4. Sequences of low tone--high tone often change so as to reduce the pitch interval between the two tones, but this happens less often with sequences of high tone--low tone (cf. Hyman 1973).

If further research supports the claim of universality for these and the many other tendencies of tonal behavior discussed in this volume, then it is clear that these patterns can best be explained by reference to universal characteristics of the human speech production and speech recognition systems.

The primary emphasis of this paper is on the physiological mechanisms involved in the production of tone; the perception of tone is not covered extensively not because it is less important, but because less is known about it (however, see Fry 1960, Hadding-Koch and Studdert-Kennedy 1964, Lehiste 1970, van Katwijk and Govaert 1967, Klatt 1973).

1. PHYSICAL CORRELATES OF TONE

Although the primary physical correlates of what has been labelled "tone" are the level, direction, and/or rate of change of the "pitch" or fundamental frequency (F_0) of vibration of the vocal cords, secondary correlates may in some cases be the mode of vibration (lax, breathy, tense, creaky voice), vowel duration and quality, and manner of termination of the vowel bearing the tone, i.e. with vs. without a glottal stop (Pike 1948, Hinton 1970, Matisoff 1973). In this paper I will be primarily concerned with the physiological mechanisms that affect F_0.

2. DETERMINANTS OF F_0

The F_0 of voice is determined basically by two partially independent factors: (a) the state of the vocal cords and (b) the aerodynamic forces driving the vocal cords (van den Berg 1958, Ishizaka and Flanagan 1972). The state of the vocal cords includes, for example, the anterior-posterior and vertical tension of the cords and the degree of approximation of both the ligamental (the anterior 2/3) and arytenoidal (the posterior 1/3) of the vocal cords. The aerodynamic driving forces are the air pressure difference ("pressure drop") across the vocal cords and the rate of air flow through the glottis.

Both of these parameters can be varied in an active, intentional way, and in a passive, fortuitous way--fortuitous in the sense that when a speaker intentionally produces some other speech gesture, these parameters may change and thus fortuitously affect F_0. (This is identical to Wang and Fillmore's (1961) use of the terms "extrinsic" and "intrinsic", respectively, with respect to phonetic variation.)

2.1. Active Variation of F_0 by Adjustment of Vocal Cord Tension

Active, intentional changes of vocal cord tension or glottal configuration are accomplished primarily by the muscles attached to the larynx, principally the cricothyroid muscles, but also by all the adductor muscles of the larynx: the lateral cricoarytenoid, the vocalis, and the interarytenoid muscles (Ohala 1970, Sawashima 1970). The extrinsic laryngeal muscles, chiefly the so-called "strap" muscles of the neck, which attach either directly to the larynx (thyrohyoid and sternothyroid muscles) or indirectly via attachment to the hyoid bone, which, in turn, is connected to the larynx via ligamental attachments (sternohyoid muscles), are also active in pitch regulation, especially in lowering pitch (Ohala and Hirose 1969, Ohala 1970, Ohala 1972, Erickson and Abramson 1972). The involvement of the strap muscles in regulating pitch is obviously related to the common observation that larynx height correlates closely with pitch (Ohala 1973, Ohala 1972, Ohala and Ewan 1973), although why this happens is still disputed. Nevertheless, I view this as evidence for the existence of active laryngeal mechanisms for both the raising of pitch and the lowering of pitch. These mechanisms are the primary ones utilized for the production of intended linguistically significant pitch changes, that is, the large pitch variations characteristic of all tonal and intonational contours. There are other views on this matter, however (Lieberman 1967, 1970).

2.2. Active Variation of Pitch by Modification of Aerodynamic Factors

The pressure drop across the vocal cords can be actively increased or decreased by varying the expiratory force. Liberman (1967) claimed that this is the primary method speakers use to increase pitch on stressed, especially emphatically stressed syllables, and to let pitch fall at the end of sentences. There are serious problems with this theory (Vanderslice 1967, Kim 1968, Ohala and Ladefoged 1969, Ohala 1970, MacNeilage 1972). Electromyographic (EMG) and other evidence points instead to the activity of the laryngeal muscles being the primary force behind these and all other major linguistic pitch variations. Moreover, recently Lieberman seems to have abandoned many of the more extravagant aspects of his theory (Lieberman, Sawashima, Harris and Gay 1970).

However, Ladefoged (1967, 1971) notes that an extra expiratory pulse (which can be detected as a momentary increase in the EMG of the internal intercostal muscles) does occur during some stressed syllables and there is a momentary increase in subglottal air pressure (P_s) during stressed syllables (which all investigators of P_s have noticed). No doubt a good part of this momentary rise in P_s is due to this expiratory pulse, but part of it may be due to a momentary increase in glottal resistance which would result from the vocal cord adjustment for increased pitch and intensity. At any rate, these P_s rises accompanying stress cannot account for most of the pitch changes observed during stress: the effect of P_s changes on pitch has been found to be about 2.5 Hz/cm H_2O (Ohala and Ladefoged 1969, Hixon, Mead and Klatt 1971), so the P_s rise of 1 to 5 cm H_2O usually found during emphatically stressed syllables could account for pitch rises of at most some 13 Hz--far less than the 50 to 100 Hz pitch changes commonly encountered on these syllables. And needless to say, the increase in P_s cannot account for any pitch drops which are commonly used to manifest stress (Bolinger 1958).

It would appear, then, that the extent to which pitch is actively regulated by variations of the expiratory force is negligible.

The aerodynamic driving force can also be modified by action of the larynx: other things being equal, the volume of air flow through the glottis will vary inversely with the glottal air resistance and the pressure drop across the vocal cords will vary inversely with the volume of air flow through the glottis. But, again, the effect on pitch of this laryngeally-induced change of the aerodynamic driving force is very likely negligible for the reasons mentioned above.

2.3. Passive Variation of Pitch by Changes in Vocal Cord Tension

2.3.1. Pull on larynx by tongue. Apparently a high position of the tongue creates a slight pull on the larynx which is translated into increased vocal cord tension. This results in the widely-noted slightly higher average pitch for high vowels [i, u] and slightly lower pitch for low vowels [æ, a, ɔ], with mid vowels having a pitch intermediate between these. This explanation is disputed, however (Atkinson 1973, but see review by Ohala 1973).

For a possible (rare) instance of this effect leading to tone alternations determined by vowel height, see Wang (1970) or Mohr (1971).

This would also predict that consonants involving a high position for the body of the tongue, i.e. palatals and velars, ought also to induce a slightly higher average pitch (as opposed to labials and dentals, which do not involve the body of the tongue in the same way). The available data on this is mixed: House and Fairbanks (1953) and Mohr (1971) do show average pitch to be higher in the environment of English [g] than it is in the environment of English [b] or [d]; however the data of Lehiste and Peterson (1961) and Lea (1972) do not.

2.3.2. Vocal cord state and voicing. Halle and Stevens (1971) suggest that the vocal cords are stiff during the production of voiceless obstruents and are slack during voiced obstruents, and they provide a model of vocal cord vibration that shows why these differences are required. This, they say, accounts for the pitch variations accompanying voiced and voiceless obstruents. However, as far as I know there has as yet been no experimental verification of their model. Recent electromyographic recordings of the intrinsic laryngeal muscles (Hirose, Lisker and Abramson 1973) revealed no obvious differences in the tension of the laryngeal muscles during the production of the voiced/voiceless distinction--other than what would be expected for the abduction of the cords during the voiceless obstruents. A crucial point in Halle and Stevens' argument is their claim that the pitch of vowels preceding obstruents also show the same kind of perturbation as do the vowels following obstruents, i.e. pitch raised slightly for voiceless obstruents; lowered slightly for voiced obstruents. But I know of no hard evidence in support of this; in fact, just the opposite: pitch usually falls immediately at the beginning of any obstruent, voiced or voiceless (Lea 1972, 1973).

2.3.3. Larynx elevation and voicing. La Raw Maran (1971) posits a raised larynx for voiceless obstruents and a lowered larynx for voiced obstruents. There is some support for this in the older literature (see Ohala 1973) and in the recent work of Ewan and Krones (1973), and given that, other things being equal, larynx height is correlated with pitch, one might expect this larynx position for consonants to affect pitch on surrounding vowels.

But perhaps "other things" are not equal in these cases because Ewan and Krones did not find a consistent relation between pitch and larynx height over the small range of larynx displacements accompanying voiced and voiceless consonants. Moreover, like the Halle and Stevens hypothesis, this would also predict that consonants would perturb pitch on both preceding and following vowels, but, as noted above, the perturbation on the preceding vowels is not the same as that on the following vowels. Thus this hypothesis certainly requires further study but there is at present no overwhelming data leading us to believe larynx height causes the pitch variations associated with voiced/voiceless obstruents.

2.3.4. Voice quality and vocal cord tension. Ohala and Ohala (1972a,b) have speculated that the lowered pitch accompanying Hindi breathy-voiced stops could be due to a lessening of the tension of the vocal cords primarily intended to achieve lowered glottal resistance in order to produce the rapid air flow through the glottis and consequently the noisy quality of breathy-voiced phonation. However, there is no evidence for this as yet.

2.4. Passive Variations of Pitch by Changes in Aerodynamic Conditions

During voiced obstruents the decrease in the pressure drop across the vocal cords (due to the build up of air pressure in the oral cavity) causes the transglottal air flow and consequently the fundamental frequency of phonation to lower (cf. Ladefoged 1967:47ff). Thus, upon release of a voiced obstruent the pitch is initially very low and may gradually return to its "normal" level. Upon release of a voiceless aspirated consonant, however, the rate of air flow is initially very high, since there is momentarily little resistance to the air flow at the open glottis or in the oral cavity. Thus when the vocal cords do adduct for voicing they meet a very high rate of air flow and consequently vibrate at an initially high rate, gradually returning to their "normal" rate of vibration.

It is interesting that pitch remains perturbed for a relatively long time after the consonant release--for as much as 100 ms after the onset of voicing (Lea 1972, Löfqvist 1973). This may be due to some kind of oscillatory momentum. Or, it is also possible that the rate of transglottal air flow characteristic of the obstruent persists for a rather long time after the stop is released. Thus air flow records published by Subtelny, Kho, McCormack and Subtelny (1969) show higher-than-normal air flow persisting more than 300 ms into the vowel following a voiceless aspirated stop. Similar records but with the high air flow lasting only about 50-60 ms into the vowel have been published by Klatt, Stevens and Mead (1968) and Frøkjaer-Jensen, Ludvigsen and Rischel (1971).

Implosives must cause elevated pitch by the high rate of air flow through the glottis generated by the rapidly descending larynx (Ladefoged, personal communication).

Breathy-voiced stops such as Hindi [b^h, d^h] etc. and voiced "h", [ɦ], are complex cases. Upon release of the breathy-voiced stops and throughout the [ɦ] the air flow is very high and thus one might expect them to raise pitch, yet they depress pitch. The reasons for this may be one or more of the following:

1. The vocal cords have to be rather closely adducted for the high rate of air flow to cause increased pitch (this is reasonable, given what we know of how the Bernoulli effect works; cf. Ladefoged 1973), and as the vocal cords are somewhat abducted during these sounds, the high air flow has little effect on pitch.
2. The vocal cord tension may be lowered in the process of abducting the vocal cords, particularly by the lateral cricoarytenoid muscles, known to play a secondary role in pitch regulation. A brief period of inhibition of the lateral cricoarytenoid muscle during [h] is evident in the EMG records presented by Ohala (1970:72).
3. Most of the air flowing through the glottis during these sounds is escaping through the arytenoidal portion, not the ligamental portion of the vocal cords (Ladefoged 1973). Thus, although the average transglottal air flow may be high during breathy-voice, the air flow through

the ligamental portion may be lower than normal. As it is only or mainly the ligamental portion of the vocal cords which is vibrating in this case this lower-than-normal air flow may cause the lowered pitch.

3. SPEED OF PITCH CHANGE

Ohala and Ewan (1973) found that for a given pitch interval subjects could execute a falling pitch faster than a rising pitch. They tentatively concluded that raising pitch may involve more "effort". This may account for (a) the aspect of "downdrift" which involves the lowering of successive high tones (Hyman and Schuh 1972, Hyman 1973), (b) the higher incidence of falling tones over rising tones in Chinese (Cheng 1973), and (c) the assymmetrical behavior of sequences of low-high tones vs. high-low tones (Hyman and Schuh 1972, Hyman 1973).

4. SYNCHRONIZATION OF TONE (ACCENT) AND SEGMENTS

There is evidence that at least in Swedish, and possibly Dutch, and very likely, English, pitch contours and segments need not be precisely synchronized, at least not as closely synchronized as the oral articulators are with each other (Öhman 1965, van Katwijk and Govaert 1967, Löfqvist 1973).

5. PERCEPTUAL RELEVANCE OF TONAL DIFFERENCES TO CONSONANT IDENTIFICATION

Although the pitch perturbations induced by different types of consonants are typically quite small, there is ample evidence that in languages not considered to be tonal (e.g. English and Russian), these pitch changes do have value as perceptual cues for the identification of the voicing of stops (Chistovich 1968, Haggard, Ambler and Callow 1970, Fujimura 1971).

If these supposedly small fortuitous pitch contours following consonants can be used as perceptual cues by listeners, it is a small step beyond that to suppose that eventually these small pitch differences might be taken by listeners as the major acoustic cue differentiating the lexical items formerly differentiated by voicing or voice onset time. Thereafter the pitch contours after the stops would be greater and would be intentionally produced. This is undoubtedly what happened over the centuries in those cases where voicing contrasts have been lost and tonal differences on the following vowel have been left in their place (cf. Maran 1973, Matisoff 1973). This is the mechanism called "rephonologization" by Jakobson (1931).

NOTE

*A somewhat puzzling exception to this generalization may be the case of the Punjabi high tones which appear on vowels which were followed by [h] or breathy voiced stops in Middle Indo-Aryan (Arun 1961, Gill and Gleason 1969).

REFERENCES

Arun, V. B. 1961. A comparative phonology of Hindi and Panjabi. Panjabi Sahitya Akademi.

Atkinson, J. E. 1973. Intrinsic F_0 in vowels: physiological correlates. (Abstract) Journal of the Acoustical Society of America 53.346.

Berg, Jw. van den. 1958. Myoelastic-aerodynamic theory of voice production. JSHR 1.227-244.

Bien-Ming, C. 1931. The tone behavior in Hagu: an experimental study. Archives Néerlandaises de Phonétique Expérimentale 6.6-45.

Bolinger, D. 1958. A theory of pitch accent in English. Word 14.109-149.

Chang, K. 1972. The reconstruction of Proto-Miao-Yao tones. Presented at the Fifth Sino-Tibetan Conference, Ann Arbor.

Cheng, C.-C. 1973. A quantitative study of Chinese tones. Journal of Linguistics 1.93-110.

Chistovich, L. A. 1969. Variations of the fundamental voice pitch as a discriminatory cue for consonants. Soviet Physics: Acoustics 14.372-378.

Erickson, D. and A. Abramson. 1973. Electromyographic study of the tones of Thai. (Abstract) Journal of the Acoustic Society of America 53.346.

Ewan, W. and R. Krones. 1973. A study of larynx height in speech using the thyroumbrometer. Abstract. Journal of the Acoustic Society of America 53.345.

Frøkjaer-Jensen, B., C. Ludvigsen and J. Rischel. 1971. A glottographic study of some Danish consonants. In Form and Substance, 123-140. Akademisk Forlag.

Fry, D. B. 1960. Linguistic theory and experimental research. Transactions of the Philological Society (London), 13-39.

Fujimura, O. 1971. Remarks on stop consonants--synthesis experiments and acoustic cues. In Form and Substance, 221-232. Akademisk Forlag.

Gill, H. S. and H. A. Gleason, Jr. 1969. A reference grammar of Punjabi. Punjabi University.

Greenberg, J. 1970. Some generalizations concerning glottalic consonants, especially implosives. International Journal of American Linguistics 36.123-145.

Hadding-Koch, K. and M. Studdert-Kennedy. 1964. An experimental study of some intonation contours. Phonetica 11.175-185.

Haggard, M., S. Ambler and M. Callow. 1970. Pitch as a voicing cue. Journal of the Acoustical Society of America 47.613-617.

Halle, M. and K. N. Stevens. 1971. A note on laryngeal features. Quarterly Progress Report No. 101.198-213. Research Laboratory of Electronics, M.I.T.

Haudricourt, A. G. 1954. De l'origine des tons en viêtnamien. Journal Asiatique 242.69-82.

Haudricourt, A. G. 1961. Bipartition et tripartition des systèmes de tons dans quelques languages d'extrême-orient. Bulletin de la Société Linguistique de Paris 56.163-180.

Hinton, B. C. 1970. Spectrographic confirmation of contrastive pitch and breathiness in Gurung. Occasional Papers of the Wolfenden Society on Tibeto-Burman Linguistics, Vol. III, Part I. Urbana, Illinois.

Hirose, H., L. Lisker and A. Abramson. 1973. Physiological aspects of certain laryngeal features in stop production. (Abstract) Journal of the Acoustical Society of America 53.294-295.

Hixon, T., J. Mead and D. Klatt. 1971. Influence of forced transglottal pressure changes on vocal fundamental frequencies. (Abstract) Journal of the Acoustical Society of America 49.105.

House, A. S. and G. Fairbanks. 1953. The influence of consonant environment upon the secondary acoustical characteristics of vowels. Journal of the Acoustic Society of America 25. 105-113.

Hyman, L. M. 1973. The role of consonant types in natural tonal assimilations (in this volume).

Hyman, L. M. and R. G. Schuh. 1972. Universals of tone rules. Working Papers in Language Universals 10.1-50. Stanford University.

Ishikawa, K. and J. L. Flanagan. 1972. Synthesis of voiced sounds from a two-mass model of the vocal cords. Bell System Technical Journal 51.1233-1268.

Jakobson, R. 1931. Principes de phonologie historiques. Travaux du Cercle Linguistique de Prague 4.247-267.

Katwijk, A. van and G. A. Govaert. 1967. Prominence as a function of the location of pitch movement. IPO Annual Progress Report 2.115-117.

Kim, C-W. 1968. Review of Lieberman (1967). Language 44.830-842.

Klatt, D. H. 1973. Discrimination of fundamental frequency contours in synthetic speech: implications for models of pitch perception. Journal of the Acoustical Society of America 53.8-16.

Klatt, D. H., K. N. Stevens and J. Mead. 1968. Studies of articulatory activity and airflow during speech. In Bouhuys, A. (ed.), Sound production in man. Annals of the New York Academy of Sciences 155.1:42-55.

Ladefoged, P. 1967. Three areas of experimental phonetics. Oxford University Press.

Ladefoged, P. 1971. Preliminaries to linguistic phonetics. University of Chicago Press.

Ladefoged, P. 1973. The features of the larynx. Journal of Phonetics 1.73-83.

Lea, W. A. 1972. Intonational cues to the constituent structure and phonemics of spoken English. Purdue Doctoral Dissertation.

Lea, W. A. 1973. Segmental and suprasegmental influences on fundamental frequency contours (in this volume).

Lehiste, I. 1970. Suprasegmentals. M.I.T. Press.

Lehiste, I. and G. Peterson. 1961. Some basic considerations in the analysis of intonation. Journal of the Acoustical Society of America 33.419-425.

Lieberman, P. 1967. Intonation, perception and language. M.I.T. Press.

Lieberman, P. 1970. A study of prosodic features. Status Report on Speech Research (Haskins Laboratories) SR-23.179-208.

Lieberman, P., M. Sawashima, K. S. Harris and T. Gay. 1970. The articulatory implementation of the breath-group and prominence. Language 46.312-327.

Löfqvist, A. 1973. On intrinsic and extrinsic F_0 variations in Swedish tonal accents. Monthly Internal Memorandum, Phonology Laboratory, Berkeley. March, pp. 19-32.

MacNeilage, P. 1972. Speech physiology. In Gilbert, J. (ed.), Speech production and perception: their relation to cortical function. New York: Academic Press.

Maran, L. R. 1971. Burmese and Jinghpo: a study of tonal linguistic processes. Occasional Papers of the Wolfenden Society on Tibeto-Burman Linguistics. Vol. IV. Urbana, Illinois.

Maran, L. R. 1973. On becoming a tone language: a Tibeto-Burman model of tonogenesis (in this volume).

Matisoff, J. A. 1973. Tonogenesis in Southeast Asia (in this volume).

Mohr, B. 1971. Intrinsic variations in the speech signal. Phonetica 23.65-93.

Ohala, J. 1970. Aspects of the control and production of speech. Working Papers in Phonetics (UCLA) No. 15.

Ohala, J. 1972. How is pitch lowered? (Abstract) Journal of the Acoustical Society of America 52.124.

Ohala, J. 1973. Explanations for the intrinsic pitch of vowels. Monthly Internal Memorandum. Phonology Laboratory. January, pp. 9-26.

Ohala, J. and H. Hirose. 1969. The function of the sternohyoid muscle in speech. Reports of the Autumn meeting (1969) of the Acoustical Society of Japan, 359-360. Reprinted in, Annual Report No. 4, Research Institute of Logopedics and Phoniatrics, University of Tokyo, Japan, June 1970, 41-44.

Ohala, J. and P. Ladefoged. 1969. Further investigation of pitch regulations in speech. Working Papers in Phonetics (UCLA) 14.12-24.

Ohala, J. and W. Ewan. 1973. Speed of pitch change. (Abstract) Journal of the Acoustical Society of America 53.345.

Ohala, M. and J. Ohala. 1972a. The problem of aspiration in Hindi phonetics. (In Hindi) Bhasha. Reprinted in English in The Annual Bulletin of the RILP, University of Tokyo, 6.39-46; and in Project on Linguistic Analysis (UCB) 63-70.

Ohala, M. and J. Ohala. 1972b. The phonetic specification of aspirates in Hindi. Paper delivered at the Second Annual California Conference on Linguistics, U.C.L.A.

Öhman, S. 1965. On the coordination of articulatory and phonatory activity in the production of Swedish tonal accents. STL-QPSR (Stockholm) 2.14-19.

Pike, K. L. 1948. Tone languages. Ann Arbor, Michigan.

Sawashima, M. 1970. Laryngeal research in experimental phonetics. Haskins Status Reports in Speech Research SR-23.69-115.

Subtelny, J. D., G. H. Kho, R. M. McCormack and J. D. Subtelny. 1969. Multidimensional analysis of bilabial stop and nasal consonants--cineradiographic and pressure-flow analysis. The Cleft Palate 6.263-289.

Vanderslice, R. 1967. Larynx vs. lungs: cricothyrometer data refuting some recent claims concerning intonation and archetypality. Working Papers in Phonetics (UCLA) 7.69-79.

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

Wang, W. S-Y. and C. J. Fillmore. 1961. Intrinsic cues and consonant perception. JSHR 4.130-136.

Wang, W. S-Y. 1970. The many uses of F_0. In Valdman, A. (ed.), Studies in linguistics and phonetics in memory of Pierre Delattre, 487-504. The Hague: Mouton.