Studies in African Linguistics
Volume 10, Number 3, December 1979

ONCE MORE ON THE NATURE OF DOWNDRIFT

Lee A. Becker
University of Illinois

In this paper is proposed a motivation for the unmarked, neutral intonation to end on a low note, and it is suggested that the most natural means of implementing this pattern may be a gradual fall, which reflects most closely the decreasing relaxation pressure generated by the passive forces of exhalation. Further, it is claimed that downdrift develops in stages: of lows only, of highs after lows, and of sequences of highs. Hombert's [1974] proposal that the development of downdrift may be blocked or stopped where it would threaten perceptual confusion is endorsed. Inasmuch as the implicational hierarchy of types of sequences which will exhibit downdrift in a given language which is predicted by the proposed development is consistent with various attested systems, the proposed development provides an explanation for the nonexistence of certain other types of systems.

0. Introduction

Hombert [1974] establishes a causal relationship between the extent of downdrift, that is whether or not it applies to sequences LLLLLL, HLHLHL, and HHHHHH, and the types of phonemic tonal contrast that exist in particular languages. For example, in Yoruba sequences of lows downdrift, but highs alternating with lows and sequences of highs do not. Hombert argues that the existence of a phonemic mid tone blocks downdrift of high tone in Yoruba since were highs to downdrift this could could lead to perceptual confusion. The relationships between the phonemic contrast and the extent of downdrift in several languages is represented in the chart given in (1), which is taken from Hombert's excellent study.

(1) H-H-H	L-L-L	H-L-H	Possible Contrasts	Languages
0	0	0	A, B, C	Dschang (dial. Bamileke)
0	0	1	A, B	NOT ATTESTED
0	1	0	A, C	Yoruba
0	1	1	A	Igbo
1	0	0	B, C	NOT ATTESTED
1	0	1	B	NOT ATTESTED
1	1	0		Impossible
1	1	1		Hausa, Shona

A = H-H vs. H-'H or H-M   B = L-L vs. L-'L   C = L-H vs. L-'H or L-M

Hombert's study reaches the following conclusion:

> "Downdrift is a natural, unmarked intonation with an ultimate but as yet unknown articulatory motivation. But this process can be blocked when it threatens to obscure a tonal contrast." [1974:178]

In addition to the obvious question of the unknown articulatory motivation, Hombert's study also raises the question of the non-attestation of languages with certain types or extents of downdrift; these can be seen in (1). It is not made clear whether this non-attestation is a result of our present meager knowledge of tonal systems, or whether some principled explanation could be proposed for the non-existence of such types of systems. The discussion below is concerned with the nature and the development of downdrift, and it will attempt to suggest answers to both these questions.

1. An Explanation for Downdrift

The intonational nature of downdrift is affirmed by Hombert, who offers Hausa as an example of the many languages which exhibit downdrift in statements but not in questions. The existence of a universal or near-universal neutral, unmarked intonation pattern has been noted by many scholars, e.g. Bolinger [1964], Lieberman [1967]. Bolinger [1964] refers to this as a "running-down pattern". In my opinion this "running-down" should be associated with a fall in subglottal pressure (Psg) which results from relaxation or lack of activity of expiratory muscles when the pressure generated by the passive forces of exhalation is below that required to drive the glottis. Further I contend that the "running-down pattern" can vary from language to language, for example, in what percentage of the utterance of intonational unit is characterized by the "running-down". I prefer to characterize this pattern as "ending on a low note".

The glottis requires a relatively constant transglottal pressure drop of 2-3 cmH_2O to phonate at a quiet speech level [Draper et al. 1959]. The transglottal pressure drop equals the subglottal pressure minus the supraglottal pressure. When supraglottal pressure remains the same changes in transglottal pressure drop are a function of Psg. Although supraglottal pressure does not always remain constant, in order to simplify the discussion, this condition will be assumed to be the case,[1] and therefore in the remainder of the discussion I will refer to Psg and not transglottal pressure.

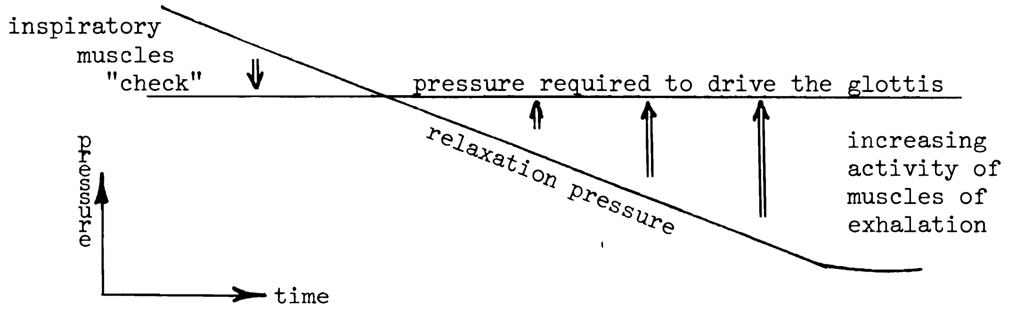
Psg is a function of the active and passive forces of respiration and the degree of aperture at the glottis. The last factor can be ignored for the discussion below (tensing the laryngeal muscles has the effect of reducing the aperture at the glottis and thus increasing Psg) since I will only be concerned with subglottally-induced changes in Psg, i.e. those that might be said to correspond to expiratory force.

The passive forces of exhalation include gravity, torque, and tissue elasticity (cf. Zemlin [1968:107] which all act to restore the lungs and rib cage to their original unexpanded and unraised state. The amount of relaxation pressure generated by these passive forces is primarily a function of the volume of air in the lungs. In (2) the Psg generated by the passive forces is represented by the diagonally-sloping line.

---

[1]In fact, it seems that expiratory force, i.e. subglottally-induced Psg, does not adjust itself to compensate for changes in supraglottal pressure in order to maintain a constant transglottal pressure drop. Consider, for example, the intrinsic amplitude of vowels. Lehiste and Peterson [1959] found that the close vowels [i u], where supraglottal pressure is greater, have a sound pressure level 5dB lower than for an open vowel like [a]; in other words, Psg does not increase to maintain the same sound pressure level. In fact perceptual tests reported by Lehiste and Peterson [1959:431] indicate that listeners equate the subjective quality of "loudness" with effort (expiratory force = subglottally-induced Psg) rather than with absolute sound pressure level. Another case would be the voiced stops where Bell-Berti [1975] has shown that rather than an increase in activity of the expiratory muscles, the larynx is lowered to increase the volume of the supraglottal cavity in order to maintain a sufficient amount of transglottal pressure drop.

(2)



The active forces are the muscles of inhalation and exhalation. When the Psg generated by the passive forces is greater than that needed to drive the glottis at a given speech (loudness) level, the muscles of inhalation contract to "check" the descent of the rib cage. According to Ladefoged [1967:25] in conversational speech at a normal level usually there is little or no checking.

In speech the muscles of exhalation become active when the relaxation pressure generated by the passive forces is insufficient to drive the larynx (at the desired level). There is increasing activity in these muscles as the relaxation pressure is further reduced, as the volume of the air in the lungs drops. However, this increasing activity in the muscles of exhalation which compensates for the continuously lowering relaxation pressure resulting from the passive forces of exhalation cannot indefinitely continue to generate the necessary Psg to drive the glottis. For one thing, the relaxation pressure may become so low, say, at the end of a very long utterance, that the contraction of the muscles of exhalation may not be able to raise the Psg to the necessary level and thus voicing may be inhibited. Second, before one can inspire again, the muscles of exhalation must relax since the Psg must be lower than atmospheric pressure for inspiration to occur. It may be suggested that this need for there to be a lower Psg to get ready for the next inspiration is a motivation for the unmarked intonation bearing a pattern of ending on a low note.

It is important to note that it is not suggested here that the fall at the end of each utterance or smaller intonational unit bearing the unmarked intonation is caused by a need to relax the expiratory muscles in order to get ready for the next inspiration, nor is it caused by the volume of air in the lungs getting so low that no amount of expiratory muscle activity will raise the Psg generated by the passive forces of exhalation enough to reach the minimum level needed to drive the glottis. The need for there to be a lower Psg would be physiologically motivated only in some cases, but the pattern of ending on a low note has been internalized and has become a general, habitual pattern.[2]

As will be suggested below different languages may generalize different strategies or means of implementation, in particular different patterns of expiratory muscle activity for achieving this low (Psg) at the end pattern. In most languages the unmarked intonation will coincide with the particular strategy, that is, the fundamental frequency curve which is used for the unmarked intonational unit, e.g. in statements, will reflect the subglottally-induced Psg, i.e. disregarding changes in Psg due to changes in the aperture at the glottis. The marked intonational unit will generally involve laryngeal modification, for example in English adjusting laryngeal tension in order to create a rise in fundamental frequency even while the Psg is falling at the end.

---

[2] Lieberman [1967] claims that the unmarked pattern of low at the end is the result of a generalization of the ontogenetically primary and innate pattern found in the cries of infants. Given that the basic cry is an attention-getting device, the infant utters a cry for as long as it can. Lieberman [1967:44] suggests that "...at the end of expiration when the lungs collapse to a certain critical point, a set of overriding respiratory reflexes automatically induce inspiration." In this connection Lieberman refers to the Hering-Breuer reflex. On my reading Widdicombe [1964:602] does not indicate that the Hering-Breuer reflex would occur in the type of situation described by Lieberman, especially in man. However, as noted above, I do believe that the need for Psg to drop in order to prepare for the next inhalation is important. Given that the pattern in question is generalized very early, it does seem to be a reasonable approach to search for a physiological motivation for this fall in the speech characteristics of infants rather than in the speech characteristics of adults; this may be the case for other aspects of language as well. It should be noted that Lieberman's [1967:104] contention that the tension of the laryngeal muscles remains steady throughout the unmarked breath group in American English has been shown to be incorrect by electromyographic studies, e.g. Hirano et al. [1969].

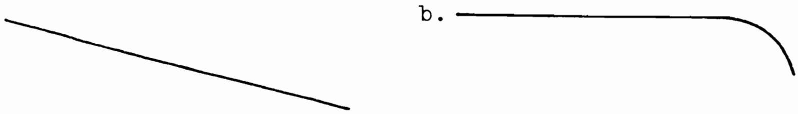
It might also be possible to speculate that the Psg pattern which most closely "shadows" or reflects the relaxation pressure (and correspondingly the volume of air in the lungs) would be the most natural. This would be a pattern which involves a lesser amount of activity of the expiratory muscles. This speculation involves the equation of less activity (of these muscles) with less effort. The principle of less effort, here involving ease of articulation, is certainly one of the primary principles of language practice and language change. It should be recalled that as the relaxation pressure continuously decreases there must be continually increasing activity of the expiratory muscles in order to maintain the same Psg.

In this section I have suggested a motivation for the unmarked, neutral intonation ending on a low note.[3] I have further speculated very tentatively that the most natural manner of implementing this fall at the end may be by a gradual fall which reflects more closely the decreasing relaxation pressure generated by the passive forces of exhalation.

2. <u>Stages in the Development of Downdrift</u>

I view the development of downdrift as having several stages. The adoption of downdrift of lows only, which I represent LLLLLL, is claimed to be the first stage. At this stage the downdrift of LLLLLL is a function of gradually decreasing Psg. In (3) are represented two patterns of Psg. (3a) might represent a prototypical downdrift pattern, while (3b) would represent a system with an extremely sharp downglide on the final syllable. Under my physiological interpretation, (3b) would reflect a relaxing of the muscles of exhalation only during the final syllable(s).

(3) a. b.



[3]Bolinger [1964] suggests that this final low corresponds to the relaxation part of a basic tension-relaxation cycle characteristic of the entire organism. The speaker tenses when he starts speaking and relaxes when he is finished (for example, after a statement as opposed to a question).

Of course, any number of intermediate stages could exist. Stage I of downdrift is a system where LLLLLL sequences exhibit a fundamental frequency pattern (corresponding to Psg) between (3a) and (3b), and where highs (in HLHLHL and HHHHHH sequences) do not exhibit downdrift.

The association of low tone with the realization of the intonational curve involves the implicit assumption that in some sense, in such a Stage I downdrift system, the low tone is the normal or neutral tone, and in languages with just an opposition between high and low, high tone represents a movement away from the normal or neutral tone. Thus I am suggesting that the reasons that lows may exhibit downdrift first is that speakers of two-tone languages may "concentrate" only on the highs and the actual pitch of the low tone, which is regarded as the absence of high tone, is of less importance. This is probably the case in many languages, but definitely not in all languages. Examples of languages where this is not the case might be the Dschang dialect of Bamileke and Nandi. As argued by Hombert [1974:176-178] Dschang does not exhibit downdrift because the presence of a downstepped low blocks downdrift and thus potential perceptual ambiguity is avoided. Creider [1978] suggests a similar motivation for the absence of downdrift in Nandi, which has a phonemic low-falling tone. I agree with Hombert's and Creider's interpretations. In Dschang, for example, the low tone would not be regarded as neutral, but would be equal to a certain percentage of a speaker's "comfortable"[4] pitch range; for a given speaker it would be equal to X cycles per second (cps). Similarly, for a given speaker other level tones would be regarded as equal to Y, Z ... cps, while if contour tones existed they might be regarded as equal to a change of E cps to F cps. In other words the tones have acoustic (or more accurately perceptual) targets, and it is a matter of whatever laryngeal tension is necessary to reach this target.

Reviewing, Stage I of downdrift is when the lows exhibit a fundamental frequency pattern (coincident with Psg pattern) between (3a), prototypical downdrift, and (3b), very abrupt downglide. A connection between

[4] This is an obvious simplification of the "normalization" problem with respect to fundamental frequency, but this question cannot be dealt with here.

downgliding and downdrift (referred to as automatic downstep[5]) has been proposed earlier by Stewart [1971]. He regarded these both as cases of "key lowering". Stewart states:

> "Generally, then, key lowering can be considered to occur at the end of every low tone syllable which is not followed by another low tone syllable; where it is non-final it is manifested as lowering of the pitch of all the subsequent high tones, and where it is final it is manifested as a slight fall towards the end of the final syllable." [1971:185]

In a footnote, Stewart suggests a causal relation between these two phenomena:

> "There is a case for taking downglide to be the more basic of the two manifestations since, if we postulate it in both contexts in sub-surface phonology, we can plausibly treat the other manifestation, namely downstep, as secondary to it; we can say that in a HLH sequence the pitch interval which separates the end of the L from the following H is the same as that which separates the beginning of the L from the preceding H, but that the following H has a lower pitch than the preceding H as the LH interval starts from the bottom of the downglide at the end of the L. The fact that in surface phonology there is no downglide in this context can be covered by a surface rule deleting downglide in non-final position." [1971:185]

Several comments must be made about this interpretation. According to Stewart's account the lows in a sequence HLHLHL would exhibit an equal amount of downdrift as would the highs in this sequence, since they would

---

[5] Stewart [1971] claims that non-automatic downstep can only arise from a system with automatic downstepping (in our terms downdrifting). In other words he always sees the development of non-automatic downstepping as the phonemicization of a previously predictable lower realization of a high after a low, as a result of loss of the low, i.e. /HLH/ [HL'H] → /H'H/ [H'H] vs. /HH/ [HH] unchanged. Although I would admit that this type of development—predictable feature becoming distinctive as a result of a neutralization—is certainly a common one, it is not the only pattern. Consider, for example, a compensatory lengthening diachronic process, e.g. $VC\{{V \atop C}\} \rightarrow \bar{V}C$, where it is only the gradual loss itself which results in the lengthening (and not any previously existing length being phonemicized when a segment is lost). It certainly seems reasonable that such a "compensatory" development (a displacing of the contrast) could result in non-automatic downstep arising from a system without automatic downstep. As the syllabic carrying the low tone is lost, its drop in laryngeal tension is in a sense taken on by the following high tone; in this manner pertinent distinctions can be maintained.

always begin a given number of cps below the preceding high. This is not the case. It is well-known, e.g. Hyman & Schuh [1974:85], Hombert [1974: 175], that lows in a sequence HLHLHL downdrift slower than the highs in such a sequence. For Stewart the downdrifting (automatic downstepping in his terms) of highs is a function of the non-surfacing downglide of non-final lows, and the lower pitch of subsequent lows is a function of the lower level of the intervening highs. Thus the lower pitch of subsequent lows in HLHLHL sequences is not in any way connected with the downdrift of sequences of lows, which would presumably have to be considered a different, disassociated phenomenon.

Under my interpretation downdrift in a sequence HLHLHL, which I will refer to as Stage II of downdrift could only develop from a system which exhibited downdrift of lows (Stage I). As argued by Hombert, this change may be blocked if it could lead to potential perceptual confusion, for example between a downdrifted high and a mid. The change of Stage I to Stage II amounts to the highs in a sense coming to ride on the intonational curve which is the level of the downdrifting lows. High tone ceases to be equal to a given percentage of a speaker's range and for an individual speaker to a fixed number of cps, as it was at Stage I; instead the fundamental frequency of high tones varies and is "relative" to that of the low pitch. It may be useful to think of high tone after low as a fixed target amount of increase in laryngeal tension. Once this has occurred, once the high tone has been disassociated from an absolute pitch for a given speaker, it is more free to be affected by the pressures of ease of articulation and the principle of least effort. As noted above, in languages with downdrift of both low and highs in a sequence HLHLHL, the highs downdrift faster than the lows in such a sequence. In this case the less effort and ease of articulation is reflected in the lesser amount of rise from low to high. It has been shown, for example by Ohala and Ewan [1972] and Sundberg [1973], that the sequence LH takes longer to produce than an HL sequence (at least in untrained singers). The lowering of a high after a low thus will make this change easier in that it will require a shorter time.[6] Alternatively it may be suggested that once high tone ceases being perceptually defined and acquires an articulatory target,[7] i.e. a fixed amount of increase in laryngeal tension, it may not make it all the way to this target (a coarticulation effect). It is important to note that it is being suggested here that this natural assimilation of the level of a high to a preceding low is generally inhibited in non-downdrifting languages, like Dschang, or in Stage I downdrifting languages because the level of a high tone in such languages is fixed at an absolute pitch for a given speaker (and for the speech community at a fixed percentage of the speaker's "comfortable" pitch range). It is only once the target is disassociated from this fixed, absolute pitch level that the natural articulatorily motivated process can become manifested more freely. As an example of a language with such a system Hombert offers Igbo, where successive lows and high alternating with lows (HLHLHL) exhibit downdrift but sequences of highs do not.

The change of a Stage II system to a Stage III system, where sequences of highs also downdrift, amounts to the establishment of a situation where all tones downdrift. Recall how in the Stage I system it was suggested that the lows were tied to, or carried by, the intonational curve. In a Stage III language all the tones are tied to this curve. This would be similar to the situation outlined by Lea [1973; 1977] for English, whereby pitch perturbations, which in English are a function of the stress of various syllables and for a Stage III downdrift language would be a function of the tones of the syllables, are superimposed upon a sentence intonation

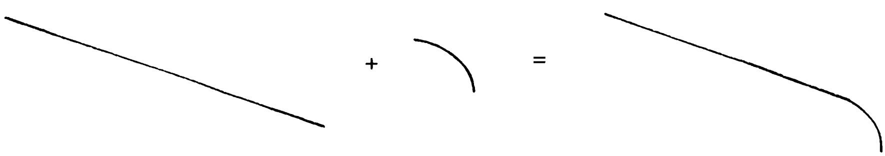
---

[6] As suggested by Hombert [1974:175], the low may possibly also be raised somewhat before a high to the same end.

[7] It should be obvious that I am claiming that in certain languages tonal features are defined in perceptual terms while in other languages they are defined in articulatory terms. My position is that any aspect or parameter of which a speaker-hearer/language-learner may be cognizant can be a potential feature, and the same "sound" can be classified, i.e. defined in terms of features, differently by (speaker-hearers of) different languages (and conceivably by speaker-hearers of the same language, especially during language acquisition).

line. One might say that in a Stage III language low tone would be associated with a certain (normal) laryngeal tension, while high tone would be associated with a certain greater (than for low) laryngeal tension; both would downdrift due to the gradually falling intonational curve, which is a function of Psg. An example of a language with a Stage III downdrift system would be Hausa, where LLLLLL, HLHLHL, and HHHHHH sequences, i.e. TTTTTT or any tone, all downdrift.

Before concluding, it should be noted that from a Stage III downdrift system where the downdrift reflects a superimposition of the tones of the syllables (in physiological terms, of certain fixed target amounts of laryngeal tension) onto the intonational pattern of the utterance (being a function of Psg), might conceivably develop a system where the gradually falling pitch of sequences of tones was dissociated from the intonation curve. In other words, the gradual fall in pitch would be no longer a function of a gradually falling Psg, but would be programmed through laryngeal tension. This development might be appropriately termed the "deintonationalization" of downdrift. From Odden's study of tone in the Karanga dialect of Shona (personal communication, 1978), it appears that such a "deintonationalization" of downdrift may have occurred there. In Karanga there appear to be two types of falling realizations of sequences of lows: one more gradual, starting from the beginning of the utterance, and a second more abrupt, occurring only at the end of the utterance when it bears a neutral intonation. This might be represented as in (4).

(4) 

The more abrupt drop looks very much like the result of a sharply falling Psg resulting from a late relaxation of the muscles of exhalation. The more gradual fall occurs not only under the unmarked intonation, but also under the marked intonation, for example as is found in conditional clauses.

3. Conclusion and Implications

Summarizing, it has been suggested that the development of downdrift is best viewed as taking place in stages. Each of the stages may be blocked, as pointed out by Hombert [1974], if potential perceptual confusion is threatened. The three stages are of lows only (LLLLLL), of highs only when separated by lows, i.e. only when after a low (HLHLHL), and of sequences of highs (HHHHHH). No positive evidence for the proposal that downdrift develops in stages, perhaps of the type as might be provided from the situations in neighboring dialects, has been offered; however, this proposed course of development does predict the following implicational hierarchy of the sequences which will exhibit downdrift:

(5) LLLLLL → HLHLHL → HHHHHH

Given that x → y, if a language has downdrift of sequence y, it is predicted that it will also have downdrift of sequence x. To the extent that this hierarchy is consistent with the facts, the proposed development is supported.

Notice that this hierarchy predicts that no languages will exhibit downdrift of HLHLHL and/or HHHHHH, but not of LLLLLL. Now look again at (1) from Hombert [1974]. Notice that three types of languages are not attested. These are systems with downdrift of HLHLHL and/or HHHHHH, but without downdrift of LLLLLL. It is, of course, possible that the non-attestation of such languages is the result of the small sample of languages listed by Hombert. The development proposed above, however, accounts for these cases of non-attestation and predicts that in a larger sampling languages with these characteristics will not be found.

REFERENCES

Bell-Berti, F. 1975. "Control of pharyngeal cavity size for English voiced and voiceless stops." *Journal of the Acoustic Society of America* 57:456-461.

Bolinger, D. 1964. "Intonation as a universal." In *Proceedings of the 9th International Congress of Linguists*, pp. 833-845. The Hague: Mouton.

Creider, C. 1978. "Nominal tone in Nandi." Paper read at 9th Conference of African Linguistics.

Draper, M.H., P. Ladefoged and D. Whitteridge. 1959. "Respiratory muscles in speech." *Journal of Speech and Hearing Research* 2:16-27.

Hirano, M., J. Ohala and W. Vennard. 1969. "The function of laryngeal muscles in regulation of fundamental frequency and intensity of phonation." *Journal of Speech and Hearing Research* 12:616-628.

Hombert, J-M. 1974. "Universals of downdrift: their phonetic basis and significance for a theory of tone." In W.R. Leben (ed.), *Papers from the Fifth Conference on African Linguistics*, pp. 169-183. *Studies in African Linguistics*, Supplement 5.

Hyman, L.M. and R.G. Schuh. 1974. "Universals of tone rules: evidence from West Africa." *Linguistic Inquiry* 5:81-115.

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

Lea, W. 1973. "Segmental and suprasegmental influences on fundamental frequency contours." In L.M. Hyman (ed.), *Consonant Types and Tone*, pp. 15-70. Southern California Occasional Papers in Linguistics, 1.

Lea, W. 1977. "Acoustic correlates of stress and juncture." In L.M. Hyman (ed.), *Studies in Stress and Accent*, pp. 83-119. Southern California Occasional Papers in Linguistics, 4.

Lehiste, I. and G.E. Peterson. 1959. "Vowel amplitude and phonemic stress in American English." *Journal of the Acoustic Society of America* 31:428-435.

Lieberman, P. 1967. *Intonation, Perception, and Language*. Research Monograph No. 38. Cambridge: MIT Press.

Odden, D. 1978. "An acoustic study of Shona tone." Paper read at LSA Summer Meeting, 1978.

Ohala, J. and W. Ewan. 1973. "Speed of pitch change [abstract]." *Journal of the Acoustic Society of America* 53:345.

Stewart, J.M. 1971. "Niger-Congo, Kwa." In T.A. Sebeok (ed.), *Linguistics in Sub-Saharan Africa*, pp. 179-212. Current Trends in Linguistics, 7. The Hague: Mouton.

Sundberg, J. 1973. "Data on maximum speed of pitch changes." *Quarterly Progress and Status Reports.* Speech Transmission Laboratory, Stockholm, 1973/74, pp. 39-47.

Widdicombe, J.G. 1964. "Respiratory reflexes." In W.O. Fenn and H. Rahn (eds.), *Handbook of Physiology, Respiration, I.* Washington, D.C.: American Physiological Society.

Zemlin, W.R. 1968. *Speech and Hearing Science - Anatomy and Physiology.* Englewood Cliffs, N.J.: Prentice-Hall.