ARIPUC 18, 1984, 1-12

GLOBAL AND LOCAL FUNDAMENTAL FREQUENCY VARIATION AND LARYNX HEIGHT: SOME PRELIMINARY OBSERVATIONS

NIELS REINHOLT PETERSEN

The paper presents data on the relation between fundamental frequency declination and larynx height, and between the stress determined fundamental frequency variation and larynx height. The results can be summarized as follows: a) Fo declination seems to be accompanied by a concomitant declination of larynx height, b) the Fo rise from stressed to the first posttonic syllable in Standard Danish is accompanied by two different patterns of larynx height variation: in one the larynx rises, in the other it lowers. The findings are discussed in the light of theories for the production of Fo.

I. INTRODUCTION

The fundamental frequency (Fo) variation in a sentence may be described as the composite result of the superposition of Fo movements associated with units of greater or lesser temporal scope, ranging from the text, through the utterance, sentence and phrase to the prosodic stress group. The segmentally conditioned Fo variation will not be considered here.

The more locally determined Fo variation is commonly agreed to be caused by the activity of the intrinsic laryngeal muscles, primarily the cricothyroid muscle (see e.g. Collier 1975, Atkinson 1978, Honda 1983, and Ohala 1978, who gives an extensive review of the literature), and to some degree assisted by the extrinsic laryngeal muscles, i.e. the inferior extrinsic muscles (the sternohyoid, sternothyroid, and thyrohyoid muscles) for Fo lowering, and the geniohyoid muscle for Fo raising. (See further Ohala 1970, Sawashima, Kakita, and Hiki 1973, Collier 1975, Erickson and Atkinson 1976, Atkinson and Erickson 1977, Erickson, Liberman, and Niimi 1977, Atkinson 1978, Honda 1983.)

The physiological mechanisms that underlie the overall fundamental frequency declination, which seems to be universally associated with terminal declarative utterances, are less agreed upon by researchers, and they have indeed been subject to less experimental research. One explanation was brought forward by Lieberman (1967), who, within his theory that claims subglottal pressure to be the major source of Fo variation, associated this type of Fo declination with the subglottal pressure declination which can be observed during such an utterance. The theory in its general form is hardly tenable (see Ohala 1978 for a review on the "larynx vs. lungs"-controversy), but the results reported in Collier 1975, where the EMG activity of a number of laryngeal muscles were examined together with subglottal pressure in Dutch sentences of varying stress patterns, indicate that the slow Fo declination could only be accounted for by subglottal pressure (whereas the local Fo variation superimposed upon it had to be - and could be - explained in terms of the activity of the laryngeal muscles examined). This explanation is repeated in Cohen, Collier, and 't Hart 1982, but recently (Collier 1983) data have been reported which suggest that it only applies under certain conditions of stress distribution in the sentence.

Another explanation - although also referring to the pulmonic system - has been advanced by Maeda (1974). On the basis of the observation that the vertical position of the larynx - like Fo - shows an overall decline over the utterance, it is hypothesized that the decreasing lung volume causes a continuous lowering of the sternum, which via ligaments and muscular tissue pulls the larynx downwards with a fundamental frequency decline as the result. Other researchers, however, have not found the larynx lowering reported by Maeda; on the contrary, Gandour and Maddieson (1976) and Ewan (1979) report that - if anything - the larynx rises during the utterance.

Maeda (1979) gives a modified version of the hypothesis. In the material presented he also fails to find the Fo decline to be accompanied by a larynx height decline[1], but he finds that the laryngeal ventricle shows an overall tendency to shorten during the utterance. Under the assumption that ventricle length can be used as an estimator of vocal fold length, the following mechanism is hypothesized: The decreasing lung volume exerts a pull upon the trachea. This pull is transferred to the cricoid cartilage and - if the thyroid cartilage is fixed - is thought to tilt the cricoid cartilage in relation to the thyroid cartilage, whereby the vocal folds are shortened and, consequently, Fo lowered.

Now, the explanations outlined above imply that gross fundamental frequency declination is an automatic consequence of inherent properties of the speech production apparatus - or more specifically, of the pulmonic system. The mere fact that Fo declination is so widespread in terminal declaratives among the languages of the world speaks in favour of this view. On the other hand, there is evidence which quite convincingly supports the notion that Fo declination is actively and purposefully controlled by the speaker: Fo declination can convey information about sentence type and function as is the case in Danish (Thorsen 1979, and 1980a) where terminal declarative sentences have the steepest declination, syntactically and lexically unmarked questions have no Fo declination, and non-final periods and interrogative sentences with word-order inversion and/or interrogative particle have intermediate degrees of declination (cp. figure 1 below). The slope of declination tends to be steeper in short than in long sentences (see e.g. Maeda 1974, Sorensen and Cooper 1980, Thorsen 1980b and 1981). This means, if declination were a by-product of the function of the pulmonic system, that the expenditure of air per unit time should be greater in short than in long utterances. This is not very probable. Resetting (partial or complete) of the declination line has been shown to take place without intervening inhalation (Sorensen and Cooper 1980, Thorsen 1980b and 1981).

The aim of the present paper is not to point at answers (let alone decisive ones) to the questions of the nature of the physiological mechanisms responsible for Fo declination, or whether declination is actively controlled or not, but merely to present a small corpus of data on grosser and finer Fo and larynx height variation, which may be taken into account in the further discussion of the questions.

II. METHOD

A. MATERIAL

The point of departure for the design of the material was Thorsen's model for Danish intonation (see e.g. Thorsen 1980a) which is reproduced in figure 1. It is seen that in declarative sentences the prosodic stress groups are superimposed upon a gradually declining line connecting the stressed syllables. The prosodic stress group in Standard Danish is constituted tonally by a relatively low stressed syllable followed by a high-falling tail of unstressed ones.

The test word used was the nonsense word [ˈfi:fi] in which the local Fo variation would be represented by the rise from the stressed syllable to the first posttonic. The variation through the sentence could be observed by inserting the test word in different places in a carrier sentence. In order to make it easier for the subjects, however, the test word was inserted in three sentences as follows:

1) *i* [ˈfi:fi] *forkortes vokalen* [i ˈfi:fi fʌˈg̊ʰɒ:d̥əs voˈg̊ʰæʔln̩]

2) *vokalen i* [ˈfi:fi] *forkortes* [voˈg̊ʰæʔln̩ i ˈfi:fi fʌˈg̊ʰɒ:d̥əs]

3) *vokalen forkortes i* [ˈfi:fi] [voˈg̊ʰæʔln̩ fʌˈg̊ʰɒ:d̥əs i ˈfi:fi].

Since the three sentences were identical with respect to stress distribution and vowel length pattern (the stressed syllables, which are underlined, were all long), it was thought justified to use this procedure and yet be able to treat the data as if they had been extracted from the first, second, and third stress group of one and the same sentence.

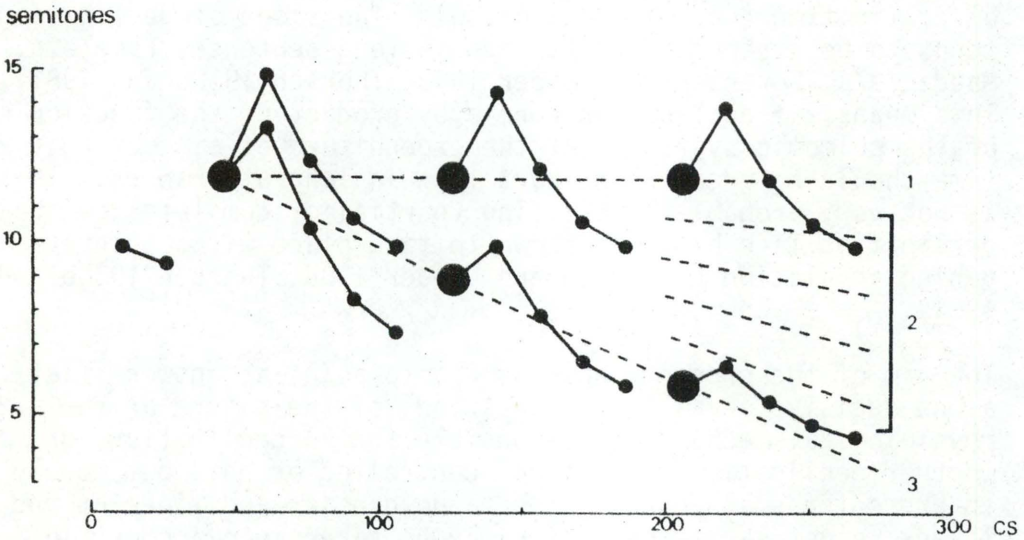

Figure 1

A model for the course of Fo in short sentences in ASC Danish. 1: syntactically unmarked questions, 2: interrogative sentences with word order inversion and/or interrogative particle and non-final periods (variable), 3: declarative sentences. The large dots represent stressed syllables, the small dots unstressed ones. The full lines represent the Fo pattern associated with stress groups, and the broken lines denote the intonation contours. Zero on the logarithmic frequency scale corresponds to 100 Hz. (Reproduced from Thorsen 1980b).

The sentences were arranged in eight random orders in a reading list.

B. RECORDING

The recording equipment consisted of a television camera (Sony AVC-3250 CES) and a video-recorder (Sony U-Matic type 2630). The frame frequency of the equipment was the normal 50 frames per second. The speech signal was recorded on the sound track of the video-tape via a Sennheiser MD 21 microphone placed about 15 cm from the subject's mouth. In order to synchronize speech and video signals a timer signal was recorded on the video-tape using a timing device (FOR-A CO. type VTG 33). On playing back the tape, the timer signal was displayed on the monitor screen in minutes, seconds, and centiseconds and it could, moreover, be registered together with the speech signal on an ink writer as pulses for seconds and centiseconds. In this manner it was possible to relate each TV-frame to the speech signal.

During recording the subject was seated in a dentist's chair with a fixed head-rest. The camera was placed at the level of the subject's thyroid prominence and at right angles to his mid-sagittal plane at a distance which allowed the area between the subject's chin and sternum to be covered by the field of vision. For calibration purposes the recording of the speech material was immediately preceded by a short recording of a millimeter scale placed in front of the subject's thyroid prominence in his mid-sagittal plane. Each subject read the list once, so that eight repetitions of each test sentence was obtained. Subjects were instructed to use a neutral declarative intonation.

C. SUBJECTS

Recordings were made of three male speakers, PD, PM, and NR (the author). PM and NR are phoneticians, and PD is an engineer and member of the technical staff of the institute. They are all speakers of Standard Danish, although PM, who has grown up in Jutland, has some dialectal influence which, however, did not seem to be reflected to any essential degree in his Fo pattern.

D. REGISTRATION AND MEASUREMENTS

The following acoustic curves were made: duplex oscillogram, two intensity curves, and an Fo curve. The timer signal was also registered on the mingograph. Measurements of Fo and larynx height were made at the midpoint of the stressed and the first posttonic vowels of the test words. The locations on the video-tape of the frames in which larynx height was to be measured were determined from the acoustic curves and the timer curve. Since the interval between frames was 2 cs (the frame frequency being 50 Hz), the temporal inaccuracy of a frame in relation to the corresponding point of measurement as determined from the acoustic curves was ± 1 cs. The video recorder was equipped with a step function which made it possible during playback to "freeze" the picture and step forward frame by frame and read off larynx height in the frames selected for measurement. Larynx height was determined from a scale drawn on the monitor screen on the basis of the millimeter scale which had been recorded on the tape prior to the reading of the material. The vertical position of the larynx could be measured with an accuracy of ± 0.5 mm.

Table I

Mean distances in cs of point of measurement from the beginning of the sentence, mean fundamental frequencies in Hz, and mean larynx heights in mm (arbitrary zero), with corresponding standard deviations, in stressed syllables ('V) and first posttonic syllables ($_{\circ}$V) in the three stress groups in the sentence.

			1st stress group		2nd stress group		3rd stress group	
			'V	$_{\circ}$V	'V	$_{\circ}$V	'V	$_{\circ}$V
NR	t	$\bar{X}$	27	47	82	102	154	177
		s	1.246	2.532	1.885	1.669	4.234	4.734
	Fo	$\bar{X}$	112	148	99	125	89	103
		s	5.685	7.285	2.843	4.999	1.769	3.464
	LH	$\bar{X}$	6.9	5.1	6.4	5.1	4.1	4.0
		s	0.835	0.885	1.061	1.246	0.641	0.0
PD	t	$\bar{X}$	21	36	66	82	116	137
		s	1.727	1.727	4.241	3.454	3.871	4.309
	Fo	$\bar{X}$	114	130	110	121	97	114
		s	5.089	5.995	9.376	7.525	2.780	4.790
	LH	$\bar{X}$	32.9	35.9	33.4	35.5	32.1	34.1
		s	1.126	0.991	0.916	1.069	0.641	0.835
PM	t	$\bar{X}$	22	39	65	82	116	138
		s	1.188	1.959	1.069	1.246	1.832	3.012
	Fo	$\bar{X}$	114	137	103	119	94	108
		s	5.731	6.776	1.750	2.840	3.382	5.894
	LH	$\bar{X}$	22.3	24.1	21.6	24.3	19.5	21.8
		s	0.463	0.991	0.916	0.707	0.926	1.581

III. RESULTS

Mean fundamental frequency, larynx height and distance from the beginning of the sentence are given in table I, and fundamental frequency and larynx height means are plotted in figure 2 as a function of time.

A. OVERALL Fo DECLINATION AND LARYNX HEIGHT VARIATION

In order to obtain a quantitative estimate of the overall Fo and larynx height variation, regression lines and correlation coefficients of these variables versus distance from the beginning of the sentence were computed. The computations were made on the basis of the raw data, and stressed and first post-

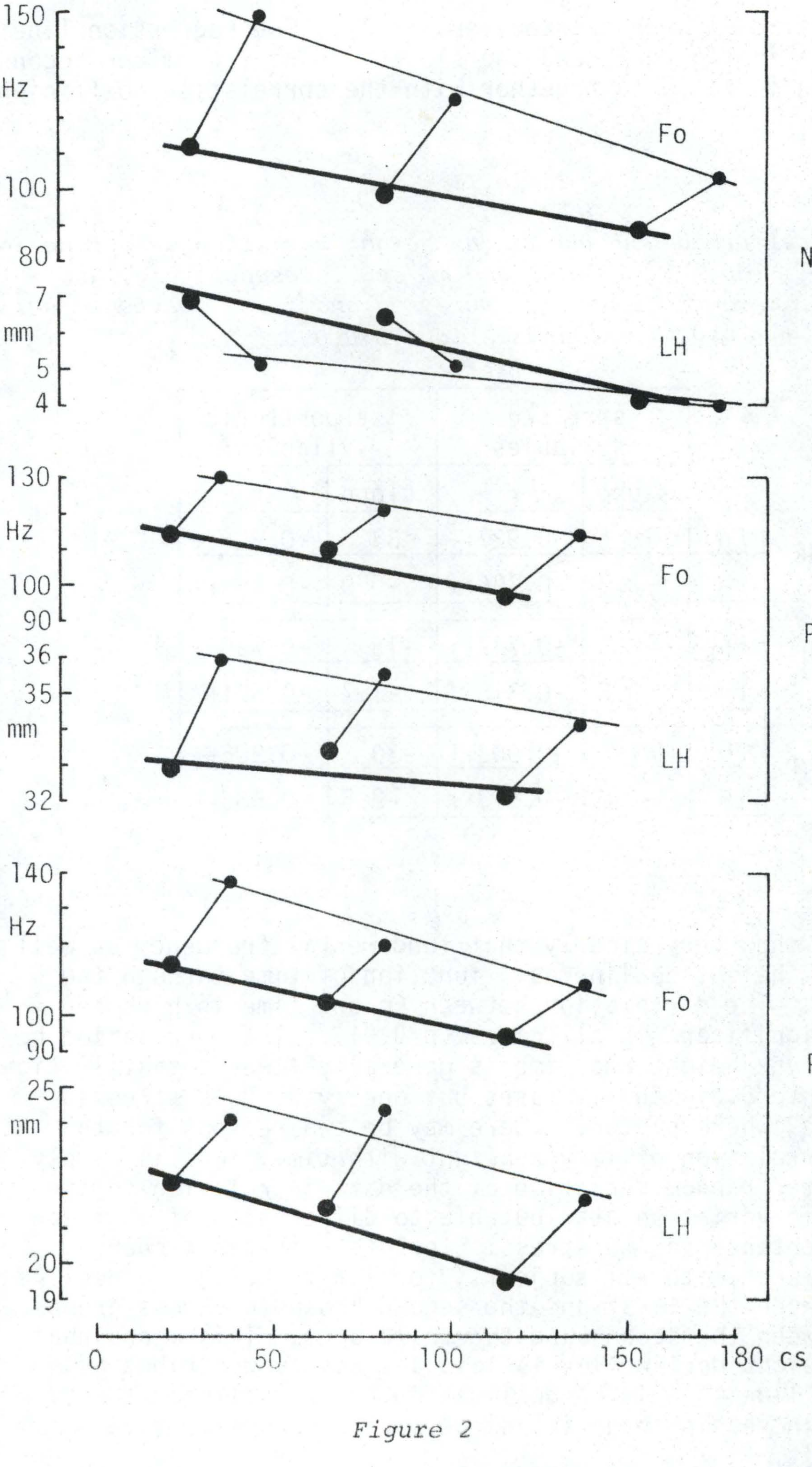

Figure 2

Fo and larynx height (in mm, arbitrary zero) as a function of time through the sentence. Stressed syllables are indicated by large and first posttonic syllables by small dots, and connected by lines within each stress group. The least squares regression lines, computed on the basis of the raw data and for stressed and first posttonic syllables separately are also shown, heavy lines for stressed syllables and light lines for first posttonics.

tonic syllables were treated separately. The regression lines are shown in figure 2, and the slopes (in Hz and mm per second) are given in table II together with the correlation coefficients.

Table II

Fundamental frequency and larynx height variation as a function of time. Slopes in Hz/sec. and mm/sec., respectively, and correlation coefficients for stressed and first posttonic syllables. ++: p<0.01, no indication: p>0.05.

		stressed syllables		1st posttonic syllables	
		slope	r	slope	r
NR	Fo	-18	-0.927++	-33	-0.955++
	LH	-2.3	-0.796++	-0.9	-0.499++
PD	Fo	-18	-0.737++	-15	-0.735++
	LH	-0.8	-0.318	-1.7	-0.601++
PM	Fo	-21	-0.902++	-30	-0.906++
	LH	-3.0	-0.813++	-2.5	-0.655++

The data show very clearly that fundamental frequency as well as larynx height decline as a function of time through the sentence. The correlation between Fo and time is high and highly significant in all cases (p<0.01). The correlation between larynx height and time is generally lower but still significant (p<0.01) in all cases but one, viz. PD's stressed syllables, where p>0.05. There may be two reasons for the lower correlation of larynx height with time. One is simply the greater random variation of the data in relation to the systematic variation attributable to differences of position in the sentence and to stress. The other is the tendency, which is seen with all subjects, for larynx height to decline less between the first and the second prosodic stress group than between the second and the third ones. This means that larynx height declination is less adequately described by a straight line than is Fo declination. A curvilinear function with an increasing negative slope would presumably give a better fit.

B. LOCAL FO- AND LARYNX HEIGHT VARIATION

Whereas there is in all cases an evident Fo rise from the stressed to the first posttonic syllable, the corresponding larynx height variation follows two different patterns (see figure 2). The two subjects PD and PM have a higher larynx position in first posttonic syllables than in stressed syllables, i.e. their larynx height pattern closely matches their Fo pattern. A one-tailed t-test showed the differences in all cases to be statistically significant ($p<0.01$) for both Fo and larynx height in stressed vs. posttonic syllable.

For the third subject, NR, the fundamental frequency is also significantly higher in first posttonic syllables than in stressed ones ($p<0.01$), but the larynx height differences go in the opposite direction, the larynx being LOWER in first posttonic than in stressed syllables. NR's pattern of larynx height variation as a function of stress seems to be less consistent than that of PD and PM. The differences are smaller and could only be proved statistically significant in the first two stress groups ($p<0.01$ and $p<0.05$, respectively).

IV. DISCUSSION

A. FO DECLINATION

The data presented above suggest that slow, overall fundamental frequency declination in terminal declaratives is somehow connected with larynx height. But the very fact that the larynx declines over the utterance makes it difficult to interpret the results in relation to the possible explanations for Fo declination outlined in section I above.

In agreement with Maeda's original hypothesis (Maeda 1974) that the decreasing lung volume lowers the sternum which in turn pulls upon the larynx via the muscular and ligamental connections between the two structures, the observed larynx lowering could be seen as the primary cause of Fo declination. Larynx lowering could also, under the subglottal pressure hypothesis, be viewed as a secondary effect of the lung volume decrease without direct influence on Fo. And, finally, if Fo declination is actively controlled, larynx lowering could be the result of activity in the inferior extrinsic laryngeal muscles, which are known to be associated with Fo lowering. Thus, the present larynx height data are compatible with any of these hypotheses.

The finding of a larynx height decline seems to weaken Maeda's (1979) tracheal pull hypothesis, since this hypothesis presupposes that the thyroid cartilage be fixed for the tracheal pull to be effective in shortening the vocal folds, which means that larynx height should remain constant over the sentence.

The possibility that overall Fo declination could be caused by a declining tension of the vocal folds as a result of a gradual decrease of overall activity in either the vocalis muscle or the cricothyroid muscle may also be considered less likely. A decrease of vocalis activity would hardly influence larynx height, and a decreasing cricothyroid activity would either have no effect on observed larynx height (i.e. the position of the thyroid prominence), if the cricoid cartilage moves as a result of cricothyroid activity, or it would raise the observed larynx position if it is the thyroid cartilage that moves.

Thus, although the observations presented in this study certainly do not make it possible to accept any single one of the hypotheses attempting to explain Fo declination and discard the others, they may to some extent narrow the field of candidates. But, of course, more research on the topic is needed, particularly research focused on the behaviour of physiological parameters under conditions which have been shown to influence Fo declination, such as sentence length and type (cp. section I above).

B. LOCAL FO VARIATION

Whereas there is good agreement among subjects with respect to the relation between overall Fo declination and larynx height, the Fo rise from stressed to first posttonic syllable is accompanied by two different patterns of larynx height variation, two subjects (PD and PM) showing an upward and the third (NR) a downward larynx movement.

Local Fo movements are primarily attributed to the activity of the intrinsic laryngeal muscles, particularly the cricothyroid, and assisted by the activity of the extrinsic laryngeal muscles. In the case of NR, his downward larynx movement may indicate that the Fo rise from stressed to first posttonic syllable is to be explained in terms of cricothyroid activity alone, since contraction of the cricothyroid muscle may cause a lowering of the observed larynx position, under the condition, of course, that cricothyroid activity rotates the thyroid cartilage and not the cricoid cartilage. If the extrinsic muscles were involved, an upward larynx movement should be expected.

The low-to-high larynx movement with subjects PD and PM can hardly be accounted for by reference to the cricothyroid muscle alone. The extrinsic laryngeal muscles will have to be taken into consideration also, and here two possibilities present themselves: either the low larynx - and Fo - of stressed syllables is brought about by contraction of the inferior extrinsic laryngeal muscles which are known to be associated with low or lowering Fo, or the high larynx - and Fo - in first posttonic syllables is due to contraction of the geniohyoid muscle, which can be active in Fo raising (cp. the references in section I above).

It seems difficult to explain the occurrence in the material of two distinctly different larynx height patterns. The conditions of recording were the same for all subjects, and they all used their normal speaking Fo range without displaying extremely large or extremely small Fo excursions. If anything, the greatest activity of the extrinsic laryngeal muscles, and hence the greatest low-to-high larynx displacement, should be expected for speaker NR, who in fact showed the opposite pattern, since he is the one who has the greatest Fo rises from stressed to first posttonic syllables.

Thus, there seems to be no other explanation than the very general one that speakers within certain limits use different strategies producing the same acoustic pattern.

V. NOTE

1. Maeda gives Fo and larynx height curves for two utterances. In one of them there is certainly no larynx height decline, but in the other there seems - to my eye - to be a larynx lowering of 6-8 mm over an utterance of about 1.5 seconds.

REFERENCES

Atkinson, J.E. 1978: "Correlation analysis of the physiological factors controlling fundamental voice frequency", *J. Acoust. Soc. Am. 63*, p. 211-222

Atkinson, J.E. and Erickson, D. 1977: "The function of strap muscles in speech: Pitch lowering or jaw opening", *Status Rep. Speech Res., Haskins Labs. 49*, p. 97-102

Cohen, A., Collier, R. and 't Hart, J. 1982: "Declination: Construct or intrinsic feature of speech pitch?", *Phonetica 39*, p. 254-273

Collier, R. 1975: "Physiological correlates of intonation patterns", *J. Acoust. Soc. Am. 58*, p. 249-255

Collier, R. 1983: "Physiological explanations for Fo declination", *Proc. 10th Int. Congr. Phon. Sc.*, p. 440

Erickson, D. and Atkinson, J.E. 1976: "The function of strap muscles in speech", *Status Rep. Speech Res., Haskins Labs. 45/46*, p. 205-210

Erickson, D., Liberman, M., and Niimi, S. 1977: "The geniohyoid and the role of the strap muscles", *Status Rep. Speech Res., Haskins Labs. 49*, p. 103-110

Ewan, W.G. 1979: "Laryngeal behavior in speech", *Rep. Phonology Lab. 3, Univ. of Calif., Berkeley*, p. 1-93

Gandour, J. and Maddieson, I. 1976: "Measuring larynx movement in Standard Thai using the cricothyrometer", *Working Papers in Phonetics, UCLA, 33*, p. 160-190

Honda, K. 1983: "Relationship between pitch control and vowel articulation", *Status Rep. Speech Res., Haskins Labs. 73*, p. 269-282

Lieberman, P. 1967: *Intonation, Perception, and Language*, (MIT Press, Cambridge, Mass.)

Maeda, S. 1974: "A characterization of fundamental frequency contours of speech", *MIT Quart. Prog. Rep. 114*, p. 193-211

Maeda, S. 1979: "On the Fo control mechanisms of the larynx", *Séminaire Larynx & Parole, Institut de Phonétique, Grenoble - 8-9 fév. 1979*, p. 245-257

Ohala, J. 1970: "Aspects of the control and production of speech", *Working Papers in Phonetics, UCLA, 15*

Ohala, J. 1978: "The production of tone", *Rep. Phonology Lab. 2, Univ. of Calif., Berkeley*, p. 63-117

Sawashima, M., Kakita, Y., and Hiki, S. 1973: "Activity of the extrinsic laryngeal muscles in relation to Japanese word accent", *Ann. Bull. Res. Inst. Logopedics and Phoniatrics, Tokyo, 7*, p. 19-25

Sorensen, J.M. and Cooper, W.E. 1980: "Syntactic coding of fundamental frequency in speech production", *Perception and Production of Fluent Speech* (ed.: Cole, R.A.) (Erlbaum Associates, Hillsdale, NJ), p. 399-440

Thorsen, N. 1979: "Interpreting raw fundamental frequency tracings of Danish", *Phonetica 36*, p. 57-78

Thorsen, N. 1980a: "A study of the perception of sentence intonation - Evidence from Danish", *J. Acoust. Soc. Am. 67*, p. 1014-1030

Thorsen, N. 1980b: "Intonation contours and stress group patterns of declarative sentences of varying length in ASC Danish", *Ann. Rep. Inst. Phon. Univ. Cph. 14*, p. 1-29

Thorsen, N. 1981: "Intonation contours and stress group patterns of declarative sentences of varying length in ASC Danish - Supplementary data", *Ann. Rep. Inst. Phon. Univ. Cph. 15*, p. 13-47