STRESS - PRESSURE CHANGES OR LARYNGEAL ACTIVITY?

Ilkka Raimo
Dept. of Phonetics
University of Turku
SF-20500 Turku
FINLAND

Olli Aaltonen
Dept. of Phonetics
University of Turku

Erkki Vilkman
Phoniatric Dept.
Tampere University
Central Hospital

ABSTRACT

We have investigated the physiological background of sentence stress production in normal and whispered Finnish, and in simulated sentences produced with excised human larynges. In normal phonation the EMG activity of the cricothyroid muscle and subglottal pressure (esophageal pressure) registered from two subjects showed a clear association with stress. In whisper the contribution of the cricothyroid muscle was negligible. The same sentence used with the living subjects was simulated by means of excised larynges. The contours could be obtained either by changing subglottal pressure only or by laryngeal adjustments only. All in all, it seems that sentence stress is not produced by any single factor but by complex interactions of physiological subsystems.

INTRODUCTION

The change in fundamental frequency is generally assumed to signal to a listener the word which has been emphasized by the speaker [1]. However, there are different views about the physiological mechanisms underlying these fundamental frequency variations. Müller [2] observed that a change in fundamental frequency can be brought about by a change in subglottal air pressure or by a change in the tension of the vocal folds. This basic distinction has been widely studied with a variety of experiments, and evidence supporting one or the other has been presented.

The essential role of the cricothyroid muscle and rotation in cricothyroid articulation in fundamental frequency (F0) regulation is well accepted (see [3] for a review). EMG studies have shown that the activity of the cricothyroid muscle correlates with fundamental frequency peaks also in sentences (e.g. [4,5,6]). Ladefoged [7] has stated that an increase in the flow of air out the lungs results an increase in fundamental frequency signalling stress. However, the subglottal pressure has been rather ineffective in fundamental frequency regulation. Such ratios as 2-5 Hz/cmH_2O has been reported in chest register [8,9]. Monsen et al. [10] have concluded that both mechanisms are involved in the production of stress, and it may be that different languages make use of these two controlling mechanisms in a different manner.

The purpose of the present study was to study the effects of the laryngeal and subglottal mechanisms on the fundamental frequency variations connected with sentence stress in Finnish. In Experiment I we studied the EMG activity of two intrinsic laryngeal muscles and esophageal pressure. The measurements were made in two different conditions: in normal phonation and in whisper. Vocal fold vibration is avoided in whisper, and the periodic voice replaced by aperiodic noise. Even with no fundamental frequency it is still possible to identify a stressed word. In Experiment II we used excised human larynges to study independently the effects of laryngeal adjustments and subglottal pressure changes on fundamental frequency signalling stress in sentence. The use of excised larynges in voice physiological studies has a long traditions. Within certain limits excised human larynges have been noticed to produce vibrations comparable to those of living subjects [2,11]. The main limitation is that the action of the thyroarytenoid muscle cannot be simulated.

MATERIAL AND METHODS

Experiment I

Two healthy, male native speakers of Finnish participated voluntarily in the investigation.

For the acoustical analysis the speech samples were recorded using a high-quality tape recorder (Teac) and an acoustic microphone in normal room acoustics. Esophageal pressure records were obtained from an air-filled (2 ml) balloon sealed to a catheter. The balloon was passed through the nose. The catheter was connected to a pressure meter (Frökjaer-Jensen Manophone). The signal from the meter was tape recorded (Racal).

EMG activities were recorded from the cricothyroid (CT) and the thyroarytenoid (VOC) muscles for normal phonation and whisper using bipolar hooked-wire electrodes. The percutaneous and submucous method was applied.

Fundamental frequency (F0) (Frökjaer-Jensen Frequency Meter) and absolute values of intensity (Frökjaer-Jensen Intensity Meter) as well as the esophageal pressure were recorded on paper. The peak value of the first, stressed syllable for each word of the three-word test utterance was measured manually.

The EMG signals were full-wave rectified and averaged (n=5) using a computer-based (Hewlett Packard) EMG data processing program developed at the Department of Biology of Physical Activity at the University of Jyväskylä [12]. The line-up point for averaging was the onset of phonation of each utterance. The peak values of EMG activity were measured from the averaged data for each word of the test utterance.

The spoken material consisted of five versions of the sentence /nalle meni ma:lle/ ("Teddy went to the country") in which the stress was varied in the following way: 1) no extra stress, 2) the first word stressed,...5) all the words stressed. The subjects repeated each version five times both for normal speech and whisper.

See [13] for further details.

Experiment II

Two normal fresh larynges taken from autopsies of males were used.

In the dissection the extralaryngeal structures were removed. Also the epiglottis and ventricular folds were dissected.

The acoustical samples were tape recorded. Electroglottographic signals were recorded using a electroglottograph (Frökjaer-Jensen EG 830) and coin-shaped electrodes attached symmetrically to the thyroid cartilage with a screw (c.f. [14]). The subglottal pressure signals were obtained using a pressure meter (Frökjaer-Jensen Manophone) connected to an outlet in the subglottal space. The pressure signal was tape-recorded (Racal). The air flow was monitored visually by means of a lead pearl flow meter attached to the air intake.

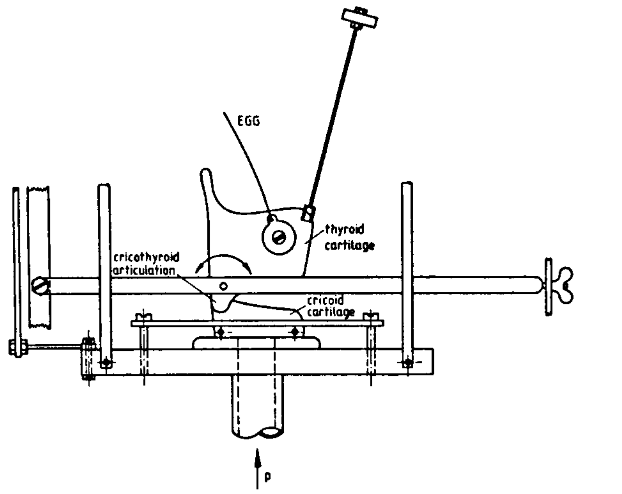

Figure 1. The apparatus used in Experiment II. EGG=electroglottograph, p=pressure. See [15] for further technical details.

The same sentences as in Experiment I were used as models and the F0 contours and the timing of this sentence were simulated to produce stress patterns 1-4.

The F0 contours were produced as follows: (1) by changing manually the laryngeal adjustments only. The flow was constant (approximately 300 ml/s). In this case the glottis was closed and opened manually and the intended stress was produced by rotating the thyroid cartilage; (2) by changing the subglottal pressure by regulating the air flow to the subglottal space. The thyroid cartilage was fixed. The glottis was closed using a forceps.

Perceptually the most natural (n=21) F0 contours were chosen by the authors and recorded on separate tape. The recordings were played to a panel of naive listeners. (n=20) whose task was to determine the position of the stress in each sentence. On the basis of this listening test the best complete sets (four sentence types) of both laryngeally and pressure induced stress patterns was formed for further analysis. The F0 (F-J Frequency Meter), intensity (F-J Intensity Meter) and subglottal pressure curves were recorded on plotting paper (Mingograph). The peak values of each variable were measured from the curves. The measuring points were the approximate site of the first stressed syllable of each word of the simulated three word sentence.

RESULTS

Experiment I

As expected, the stressed words differed acoustically from the non-stressed words in normal speech by prominent peaks in the F0 and intensity curves. In whisper stress was signalled by intensity. Each word purposely stressed by the speakers, was accompanied by a peak in the intensity curve. However, intensity did not show the same downward slope in whisper as in normal phonation. For instance, the last word in an utterance without any extra stress could exhibit the highest intensity peak.

The peak values of the esophageal pressure (p_{oes}) were nearly equal for both normal phonation and whisper. The most striking difference between them was that the differences between the peak values of the p_{oes} for the stressed and the non-stressed words were greater in whisper than in normal phonation. It is also worth noting that the pressure peaks increased towards the end of the utterances in whisper when all the words were stressed. The correlations of the p_{oes} with the acoustic variables of stress proved to be significant for both normal speech and whisper.

The cricothyroid muscle (CT) showed a peak of activation immediately preceding the peaks in normal speech. The correlations of the CT activity with F0, intensity and p_{oes} were also significant. In whisper the correlation of CT and intensity was non-significant. There were recordings in which the peaks in the CT muscle activity did not exceed the level of rest discharge for stressed words, even though they occasionally showed a phasal relationship with stress. In general, the average activation pattern of the cricothyroid muscle was lower and less variable in whisper than in normal speech.

Contrary to CT, the thyroarytenoid muscle (VOC) exhibited more similar stress patterns of the EMG activity for normal phonation and whisper. However, the two conditions differed from each other as to the correlations of the VOC-EMG with p_{oes} and intensity. The correlations between the VOC-EMG and intensity as well as between the VOC-EMG and p_{oes} were lower in whisper than in normal phonation.

The activity of the thyroarytenoid muscle also showed interindividual differences. There was a significant correlation between F0 and VOC-EMG only for one speaker. For both subjects the VOC-EMG correlated significantly with intensity. See [13] for further details.

Experiment II

The basic flow level in the laryngeally induced stress was lower than for the flow-induced stress. Consequently, intensity and subglottal pressure are higher in the latter case. Higher flow rate (approximately 500 ml/s) was needed for instance for getting an abrupt enough attack. In both cases, however, the phonation type represented chest register phonation and also the higher pressure values were within physiological limits (Fig. 2).

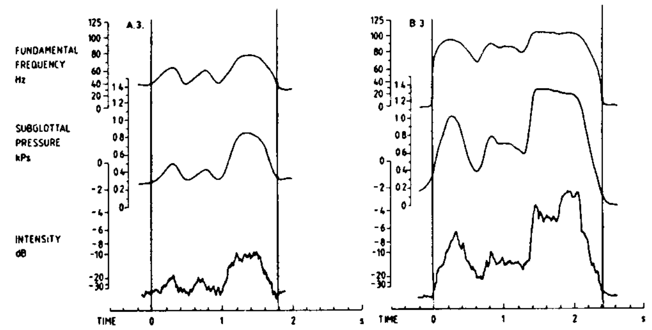

Figure 2. An example of the F0 contours produced by an excised larynx. The last word of the simulated sentence /nalle meni ma:lle/ has been emphasized. The left hand panel shows curves obtained by changing the laryngeal adjustments. The right hand panel illustrates results of changing the flow.

The differences between the peak values of the variables within each sentence for both experimental conditions were calculated. The changes in the fundamental frequency are greater in laryngeal stress than in pressure induced stress even if the subglottal pressure changes in the latter case were somewhat larger. The ratio of fundamental frequency and subglottal pressure change was 63.1 Hz/kPs for laryngeally induced stress and 26.6 Hz/kPs in the stress induced by air-flow changes.

The listeners perceived stress quite reliably irrespective of whether it was produced by changing the laryngeal adjustments only or by changing pressure only. The correlations between the peak values of F0, intensity and subglottal pressure were significant in both conditions. The intercorrelations were lower for laryngeally induced stress.

DISCUSSION

The peak values of the acoustical variables, and the cricothyroid muscle EMG activity exhibited a significant correlation. The thyroarytenoid muscle showed significant correlation with F0 only for one subject. This may be partly due to individual differences in strategies of producing stress and to technical difficulties in obtaining EMG data free of movement artifacts. The contribution of the thyroarytenoid muscle in stress production in normal speech has qualitatively been reported before (e.g. [4]). However, based on correlation analysis of the EMG data of one subject, Atkinson [6] stated that the thyroarytenoid muscle plays a minor role in controlling fundamental frequency in speech.

The minor role of the cricothyroid muscle in whisper is supported by an X-ray study of the larynx according to which the distance between the anterior tips of the cricoid and the thyroid cartilage (an estimate of the vocal fold length) was the same in whisper as in respiration [16]. The low level of cricothyroid activity in whisper when there is a need for higher effort level for stress production is comprehensible when the function of this muscle is considered: the vocal folds are abducted by cricothyroid muscle twitch [17].

Warren [18] observed higher intraoral pressures in whisper than in normal phonation. If the esophageal pressure used in the present study is a gross estimate of tracheal pressure, then our esophageal pressure peak values were only occasionally higher in whisper than in normal speech. Still, the peak values of the subglottal pressure in whispered stress production are high enough to elicit vibrations of the vocal folds in normal phonation. Biomechanically the increased thyroarytenoid muscle activity in whisper may cause the necessary extra medial compression or internal stiffening needed to hinder the vocal folds from vibrating during whispered stress production.

The correlation of the intensity of whisper and the EMG activity of the thyroarytenoid muscle was positive but nonsignificant. VOC EMG maxima, however, often appeared in the vicinity of the whispered stressed word. The correlation of the peak values of the esophageal pressure with the intensity of whisper was significant. It can be assumed that the thyroarytenoid muscle activity together with other laryngeal muscles (e.g. the lateral cricoarytenoid) and the actively increased subglottal pressure reflect the extra effort needed to cause perceivable stress in whisper. Correspondingly the low subglottal pressure values for unstressed words, as compared to normal phonation, may reflect low glottal resistance and respiratory activity.

The results of the experiments with the excised larynges are well in line with earlier findings. The increase in glottal resistance due to cricothyroid muscle twitch has been reported before in living subjects [19] and with excised larynges (e.g. [15]). The ratio of fundamental frequency and subglottal pressure change was also in the limits of earlier findings concerning chest voice phonation (e.g. [9]). However, the mechanism through which the subglottal pressure affects the fundamental frequency is not clear. It has been set forth that even in this case the reason would be laryngeal [20].

From the phonetic point of view the most important result is that perceivable sentence stress can be produced with excised larynges both by means of laryngeal and pressure adjustments. It has been suggested that the results of studies on voice production should be interpreted in terms of systems physiology because there are considerable interactions between the contributing subcomponents (e.g. [21]). All in all it seems that the dichotomic question: stress - vocal fold tension or subglottal pressure? is not justified on the basis of the present study.

REFERENCES

[1] Lehiste I: Suprasegmentals. MIT Press, Cambridge 1970

[2] Müller J: Handbuch der Physiologie des Menschen.Teil I. Hölscher, Coblenz 1837

[3] Hollien H, Hicks J W: Mechanisms for the control of vocal frequency. In: Speech Communication Papers Presented at the 97th meeting of Acoust Soc Am 12-16 June 1979. pp. 97-100. Eds. J J Wolf, D H Klatt. MIT Press, Cambridge 1979

[4] Hirano M, Ohala J, Vennard W: The function of laryngeal muscles in regulating fundamental frequency and intensity of phonation. J Speech Hear Res 12: 616-628, 1969

[5] Collier R: Physiological correlates of intonation patterns. J Acoust Soc Am 58: 249-255, 1975
[6] Atkinson J: Correlation analysis of the physiological factors controlling fundamental voicefrequency. J Acoust Soc Am 63:211-222, 1978
[7] Ladefoged P: A course in phonetics. Harcourt, Brace, Jovanovitch, New York 1975
[8] Hixon T J, Klatt D H, Mead J: Influence of forced transglottal pressure on fundamental frequency. J Acoust Soc Am 49: 105, 1971
[9] Baer T: Reflex activation of laryngeal muscles by sudden induced subglottal pressure changes. J Acoust Soc Am 65: 1271-1275, 1979
[10] Monsen R B, Engebretson A M, Vemula N R: Indirect assessment of the contribution of subglottal pressure and vocal-fold tension to changes of fundamental frequency in English. J Acoust Soc Am 64: 65-80, 1978
[11] van den Berg J W, Tan T S: Results of experiments with human larynges. Pract Oto-rhino-laryng 21: 425-450, 1959
[12] Viitasalo J. Komi PV: Signal characteristics of EMG with special reference to reproducibility of measurements. Acta Physiol Scand 93: 531-539, 1975
[13] Vilkman E, Aaltonen O, Raimo I, Ignatius J, Komi PV: On stress production in whispered Finnish. J Phonetics. In press.
[14] Lecluse FLE: Elektroglottografie. Dissertation. Erasmus University, Rotterdam 1977
[15] Vilkman E: An apparatus for studying the role of the cricothyroid articulation in voice production of excised human larynges. Folia Phoniat. In press.
[16] Fink B, Demarest R: Laryngeal biomechanics. Harvard University Press, Cambridge 1978
[17] Arnold GE: Physiology and pathology of the cricothyroid muscle. Laryngoscope 71: 687-753, 1961
[18] Warren D W: Aerodynamics of speech production. In: Contemporary issues in experimental phonetics. pp. 105-137. Ed. N J Lass. Academic Press, New York 1976
[19] Yanagihara N, von Leden H: The cricothyroid muscle during phonation. Ann Otol Rhinol Laryngol 75: 987-1007, 1968
[20] Titze I, Talkin D: A theoretical study of the effects of various laryngeal configurations on the acoustics of phonation. J Acoust Soc Am 66: 60-74, 1979
[21] Müller E M, Abbs J H, Kennedy J G: Some systems physiology considerations for vocal control. In: Vocal fold physiology. pp. 209-227. Ed. K N Stevens, M Hirano. University of Tokyo Press, Tokyo 1981

ACKNOWLEDGEMENTS

The study was financially supported by a grant from the Kordelin Foundation. The authors wish to express their gratitude to Professor Jaakko Lehtonen, Professor Paavo V. Komi and MD Jaakko Ignatius for technical and practical support.