Physiological Explanations of F_0 Declination

R. Collier[1] and C.E. Gelfer[2]
[1]*Antwerp, Belgium; Eindhoven, the Netherlands and* [2]*New York; New Haven, U.S.A.*

Declination is the tendency of F_0 contours to exhibit a tilted overall pattern: the major rises and falls appear to be superimposed on an imaginary baseline that drifts down over the entire course of the utterance. This baseline becomes actually visible during longer stretches of speech in which no major F_0 changes occur, for instance, when there is only one pitch accent in a fairly long utterance.

Cohen, Collier and 't Hart (1982) point out that the notion of declination has developed from an operational construct, useful in the interpretation of F_0 recordings, to a theoretical concept of phonetic and linguistic importance. Since declination is now being considered an intrinsic feature of speech pitch, the question regarding its physiological origin and, hence, its programming and control, becomes increasingly relevant.

1. A simple model

A simple model of pitch control in speech may take the following form:

(a) all consciously intended pitch rises and falls are effected by appropriate variations in the tension of the vocal folds; this tension is regulated by the laryngeal muscles (mainly the cricotheroid muscle and some strap muscles of the neck, such as the sternohyoid or the thyrohyoid).

(b) the gradual downdrift of the overall pitch level, i.e. declination, is caused by slowly decreasing subglottal pressure over the course of the utterance.

Part (a) of this model is well supported by quite a number of physiological experiments, which also indicate that short term variations in subglottal pressure do not significantly contribute to the production of momentary pitch inflections (see survey in Atkinson 1978 and Ohala 1978). However, these experiments are not directly relevant to part (b) of the model, which concerns long term variations in the pitch parameter. Only a subset of the data presented in Collier (1975) suggest that part (b) of the model may be a plausible account of declination. Therefore an experiment was set up in which declination could be studied in greater detail.

2. Experimental procedure

In order to observe declination in its pristine form of appearance, utterances were constructed containing no more than two pitch accents, implemented by rise-falls, while the rest of their contours shows only declination pitch (see the stylized F_0 contours in Table I). The length of the utterances, more particularly the length of the actual stretch of declination, was varied in five steps: from 5 to 18 syllables (roughly 0.7 to 3 seconds). In order to factor out the effects of segmental perturbations, the (Dutch) utterances were also mimicked in reiterant speech with /ma/ and /fa/ syllables. In all there were 45 utterance types, each read five times in succession by one subject, the first author.

Simultaneous recordings were made of, among others, the following physiological variables: subglottal pressure (P_s), recorded directly through a tracheal puncture above the first tracheal ring, and the electromyographic activity in the crico-thyroid (CT) and sternohyoid (SH) muscles, recorded with hooked-wire electrodes. The sampling and processing techniques for these parameters have been described by Harris (1981).

F_0 was measured with the algorithm designed by Duifhuis, Willems and Sluyter (1982).

3. Results

3.1. F_0 declination

After having established in a sample of data that the inter-token variability was negligibly small, only one exemplar of the five repetitions of each utterance type was selected for F_0 analysis. The F_0 values, measured at the points indicated in Table I, were similar across the conditions of length variations and normal versus reiterant speech. Therefore the F_0 values at these points were averaged and only the three accent conditions were kept separate.

Since the declination stretches vary in length, a relatively fixed ΔF_0 over a variably amount of time leads to systematic differences in the declination rate, as can be seen in Table I. This variable rate has been observed before, among others by 't Hart (1979).

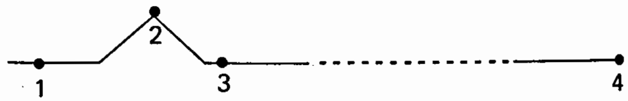
3.2. P_s declination

The P_s values, measured at the same points as the F_0 values, and averaged over all five repetitions of each utterance type, exhibit the same tendencies as the F_0 data (see Table I): they are very stable across conditions and therefore pooled together, keeping only the three accent conditions apart. It is clear that the rate of P_s declination varies with utterance length the same way as the F_0 declination does. This gross correspondence between the two variables

Table I.

A. Early pitch accent

Stylized F_0 contour and measuring points

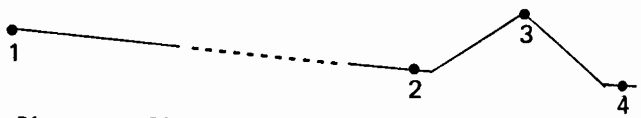


Average values (and standard deviations) at point P_i	*P1*	*P2*	*P3*	*P4*
F_0 (Hz) n = 15	109 (7)	150 (14)	99 (6)	84 (4)
P_s (cm aq) n = 75	8.3 (0.4)	9.2 (0.3)	7.7 (0.8)	4.1 (0.4)

Declination rate in length category L_i	*L1*	*L2*	*L3*	*L4*	*L5*
F_0 (Hz/sec)	-14	-12	-9	-8	-7
P_s (cm aq/sec)	-3.5	-3	-2.7	-1.9	-1.4

B. Late pitch accent

Stylized F_0 contour and measuring points

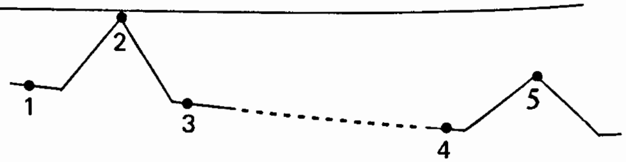


Average values (and standard deviations) at point P_i	*P1*	*P2*	*P3*	*P4*
F_0 (Hz) n = 15	115 (3)	90 (4)	117 (3)	77 (3)
P_s (cm aq) n = 75	8.4 (0.8)	5 (0.8)	7 (0.8)	2.6 (0.8)

Declination rate in length category L_i	*L1*	*L2*	*L3*	*L4*	*L5*
F_0 (Hz/sec)	-46	-23	-16	-12	-12
P_s (cm aq/sec)	-4.3	-3	-1.8	-1.4	-1.3

C. Double pitch accent

Stylized F_0 contour and measuring points



Average values (and standard deviations) at point P_i	*P1*	*P2*	*P3*	*P4*	*P5*
F_0 (Hz) n = 15	111 (5)	154 (12)	109 (6)	91 (6)	120 (4)
P_s (cm aq) n = 75	8.4 (1.5)	10.7 (0.9)	7.9 (0.9)	5.3 (0.8)	7.3 (0.6)

Declination rate in length category L_i	*L1*	*L2*	*L3*	*L4*	*L5*
F_0 (Hz/sec)	-75	-30	-11	-11	-7
P_s (cm aq/sec)	-6.4	-3	-1.9	-1.5	-1.2

suggest that F_0 declination may indeed be caused by the gradual decrease of P_s.

3.3. A causal relationship?

The extent to which P_s variations can effect F_0 changes has been studied mostly in a number of 'push in the stomach' experiments (see a survey in Bear 1979). In the chest register the P_s/F_0 ratio appears to vary between 1/3 and 1/7, which means that a ΔP_s of 1 cm aq results in a ΔF_0 of 3 to 7 Hz. Therefore a necessary condition for P_s declination to be the (sole) cause of F_0 downdrift is, that their ratio remain within these established limits.

In the 'early accent' condition of our data his is invariably the case : the P_s/F_0 ratio varies exactly between 1/3 and 1/7 in the 15 utterance types that exemplify this condition. The average ratio is 1/4.

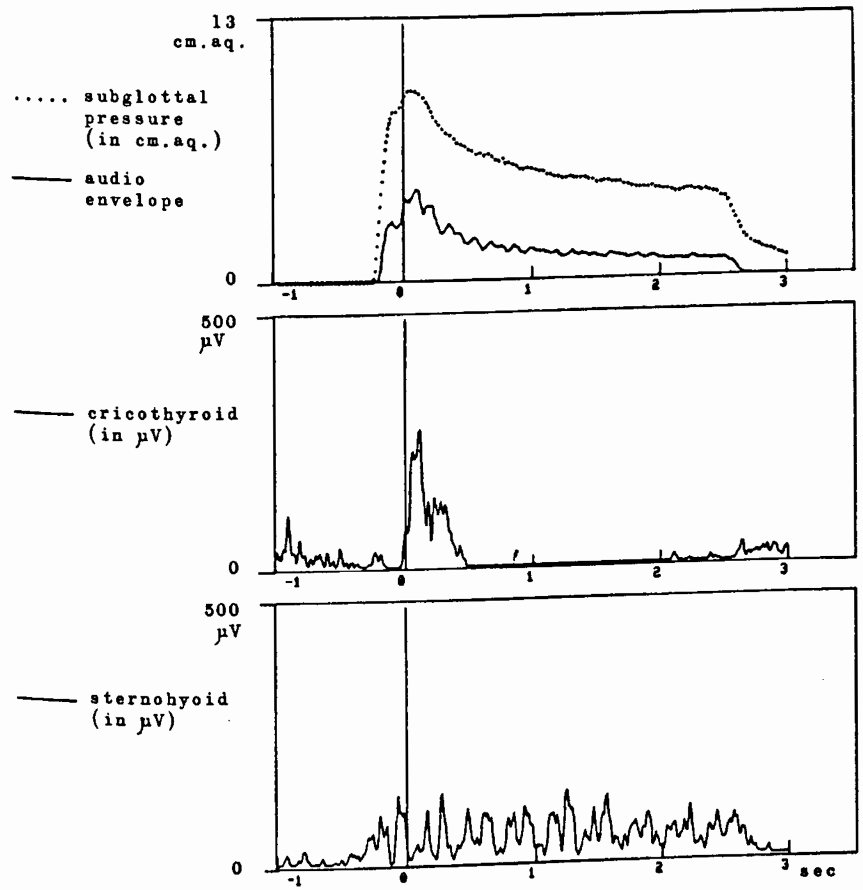


Fig. 1a. P_s and EMG data, averaged over 5 repetitions of a reiterant speech utterance with /ma/ syllables. 'Early' pitch accent at line-up point 0.

In the 'double accent' case, only 9 out of 12 utterance types comply with the established ratio. Overall the ratio varies between 1/4 and 1/11, with an average of 1/7.

In the 'late accent' condition no more than 2 utterance types have a P_s/F_0 ratio greater than 1/7. For the ensemble of this condition the ratio ranges between 1/6 and 1/16, with an average of 1/11.

Clearly, the P_s/F_0 ratio is significantly different in the three accent conditions. In the 'early accent' situation P_s declination can in itself explain F_0 declination. In the other two accent conditions the gradual decrease of P_s is often too small to account for the full extent of the F_0 downdrift.

3.4. Other factors?

In the 'early accent' condition, illustrated in Figure 1a, there is typically no

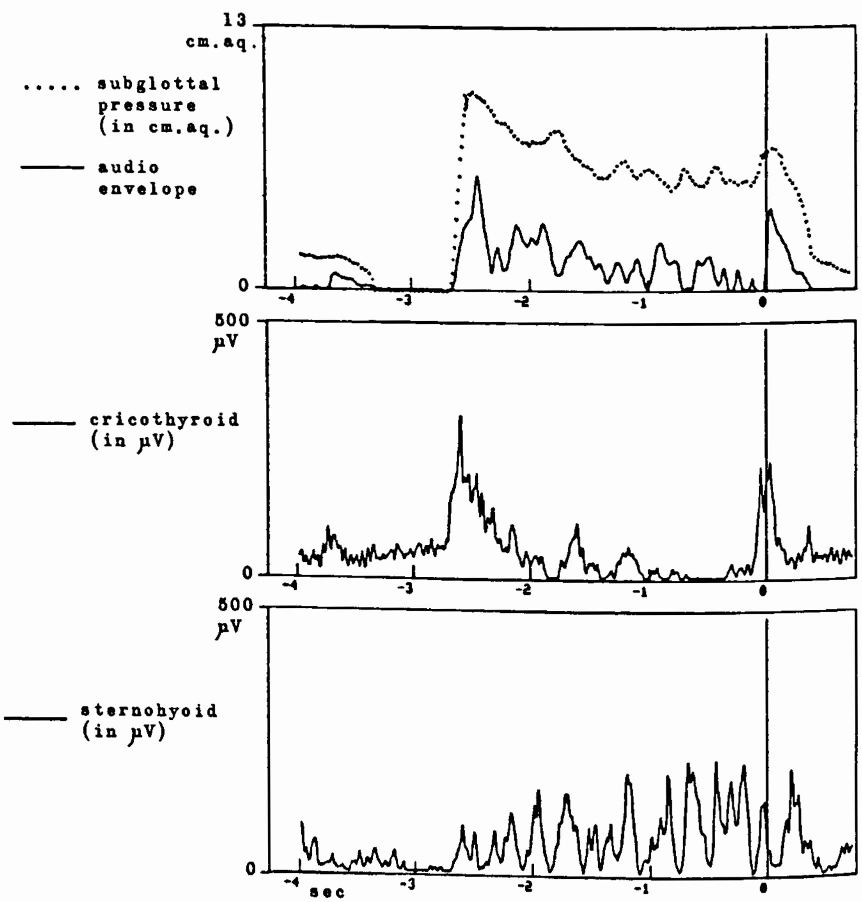


Fig. 1b. P_s and EMG data, averaged over 5 repetitions of the Dutch utterance 'Je weet dat Jan erover nadenkt ons hiervoor met genoegen te betalen'. 'Late' pitch accent at line-up point 0.

CT activity during the declination stretch and SH shows nearly equal peaks of activity, mainly associated with segmental speech gestures such as jaw lowering. In the 'late accent' condition F_0 starts at a relatively high level and this is preceded by a fairly large amount of CT activity. In many instances CT relaxes gradually (over a period of up to one second) and its relaxation is then sometimes accompanied by an increasing amount of SH contraction (see Figure 1b). Thus, the combined patterns of activity in these two muscles may account for some fraction of the F_0 lowering, in cooperation with decreasing P_s. However, this picture of combined laryngeal and respiratory action does not emerge systematically enough to explain F_0 declination whenever P_s alone cannot account for it. Moreover, the same pattern also emerges in some of the utterance types in which declination can in principle be explained by references to P_s only.

Roughly the same state of affairs holds for the 'double accent' condition.

4. Discussion and conclusion

Evidently, part (b) of our simple model accounts for the situation in which the last or only rising-falling accent occurs early in the utterance. In such a case there is no interaction of P_s with the (inactive) CT muscle or with SH (which then shows no pitch related activity). But whenever CT and (sometimes) SH are involved in pitch control, their activity is not limited to bringing about momentary pitch inflections; it can also assist in the continuous gradual pitch lowering, called declination. This means that, in certain cases, F_0 declination is not the mere byproduct of respiratory regulation but is partly controlled by laryngeal action. However our data do not show any transparant trading relationship between P_s variation and CT or SH activity. Therefore, other muscular or mechanical factors that affect vocal fold tension and glottal resistance may also be involved.

Acknowledgement

The physiological data were sampled and processed at Haskins Laboratories (New Haven, CT, USA). The F_0 measurements were made at IPO (Eindhoven, the Netherlands). We thank both institutions for their hospitality.

References

Atkinson, J.E. (1978). Correlation analysis of the physiological features controlling fundamental voice frequency. *J. Acoust. Soc. Am.* **63**, 211-222.

Baer, T. (1979). Reflex activation of laryngeal muscles by sudden induced subglottal pressure changes. *J. Acoust. Soc. Am.* **65**, 1271-1275.

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

Collier, R. (1975). Physiological correlates of intonation patterns. *J. Acoust. Soc. Am.* **58**, 249-255.

Duifhuis, H., Willems, L.F. and Sluyter, R.J. (1982). Measurement of pitch in speech: an implementation of Goldstein's theory of pitch perception. *J. Acoust. Soc. Am.* **71**, 1568-1580.

Harris, K.S. (1981). Electromyography as a technique for laryngeal investigation. In: C.L. Ludlow and M.O. Hart (Eds.), *ASHA Reports: Proceedings of the Conference on the Assessment of Vocal Pathology*, 1981.

Ohala, J.J. (1978). Production of tone. In: V. Fromkin (Ed.), *Tone: a linguistic survey*. New York, Academic Press.

't Hart, J. (1979). Explorations in automatic stylization of F_0 contours. *IPO Annual Progress Report* **14**, 61-65.