Declination : construct or intrinsic feature of speech Pitch?

Citation for published version (APA):

Cohen, A., Collier, R. P. G., & Hart, J. (1982). Declination : construct or intrinsic feature of speech Pitch? *Phonetica*, 39(4/5), 254-273. <https://doi.org/10.1159/000261666>

DOI:

[10.1159/000261666](https://doi.org/10.1159/000261666)

Document status and date:

Published: 01/01/1982

Document Version:

Publisher's PDF, also known as Version of Record (includes final page, issue and volume numbers)

Please check the document version of this publication:

- A submitted manuscript is the version of the article upon submission and before peer-review. There can be important differences between the submitted version and the official published version of record. People interested in the research are advised to contact the author for the final version of the publication, or visit the DOI to the publisher's website.
- The final author version and the galley proof are versions of the publication after peer review.
- The final published version features the final layout of the paper including the volume, issue and page numbers.

[Link to publication](#)

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal.

If the publication is distributed under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license above, please follow below link for the End User Agreement:

www.tue.nl/taverne

Take down policy

If you believe that this document breaches copyright please contact us at:

openaccess@tue.nl

providing details and we will investigate your claim.

Declination : Construct or Intrinsic Feature of Speech Pitch?

A. COHEN, R. COLLIER, J. 'T HART

Institute for Perception Research, Eindhoven, The Netherlands

Abstract. Declination is taken as the focus of studying pitch phenomena from an acoustic, physiological and perceptual point of view. It is shown that originally declination was no more than a theoretical construct to account for the interpretation of acoustic F_0 recordings. Recently, psycholinguistic considerations have enhanced the domain of application so as to account for this phenomenon. The literature is reviewed and the authors take issue over the various claims put forward by others, such as the dominance of the topline over the baseline approach, and the amount of pre-programming involved in declination, as manifested in its slope and in linguistically determined resetting.

0. Introduction

We would like to take this opportunity, afforded by the initiative of the editor of this journal, to present our views on a phenomenon that has recently become a lively issue of debate in the literature on intonation. The phenomenon to which we want to refer, declination, was first reported by PIKE [1945, p. 77]: 'The general tendency of the voice is to begin on a moderate pitch and lower the medium pitch line during the sentence. A long 'level' contour, therefore, might gradually have its syllables pronounced on a lower pitch; this can be called DRIFT.' The notion of a tilted overall F_0 pattern can also be observed in descriptions of tone languages, where this phenomenon is referred to as 'downdrift'.

The name 'declination' for this effect was coined by COHEN and 'T HART [1965]. They incorporated declination as a useful concept in the melodic model of Dutch intonation which they were developing through the following approach. The starting point was the question what the perceptual relevance was of the various F_0 fluctuations that could be observed in an acoustic recording. Only those F_0 changes that

gave rise to clearly noticeable perceptual variations were taken to be essential cues in the melodic pattern. From the outset this approach was calculated to chart the various characteristic pitch movements without, at first, bothering about specific syntactic and/or semantic features that could be signalled by them. In other words, distinctiveness in the linguistic sense was subordinated to distinctiveness in the perceptual domain.

The predominance of the perceptual point of view does not necessarily amount to the exclusive use of impressionistic auditory criteria. It implies a constant interaction between the physical and perceptual levels of description. Indeed, studying the acoustic recordings of F_0 , though in itself a major improvement on impressionistic descriptions, is by no means sufficient. The measured F_0 variations should be related to their perceptual effect and to their physiological origin.

The approach was implemented in a so-called stylization technique, applied by means of an intonator. This is in essence an F_0 generator of an analysis-synthesis system as used in vocoders. The device makes it possible to replace the natural course of F_0 in any particular utterance by a stylized artificial contour consisting of the minimal set of pitch movements that are necessary and sufficient to preserve its melodic identity. In this way one can sort out which F_0 fluctuations are perceptually relevant and which are not. This perceptual detour is indispensable because the visual conspicuousness of an F_0 change observed in an acoustic recording is not always a good measure for its perceptual salience.

Whenever an utterance stretch contains no conspicuous F_0 changes the problem arises what shape the artificial contour should be given during such a stretch. In other words, what is the natural reference pitch level from which the perceptually relevant movements depart? Our early attempts to interpret pitch movements as deviations from a constructed average horizontal, constituting the natural position of a speaker's voice setting, proved to be untenable. Closer inspection of F_0 curves revealed that a general gradual tapering off from beginning towards the end could be observed. Following these observations the natural reference level of a contour was taken to be a gradually declining baseline. This baseline and a parallel raised declination line were introduced in the stylized contours during stretches containing no major pitch movements. In subsequent listening tests the perceptual relevance of the declination effect was clearly established.

In summary, declination arose clearly as a construct which made it possible to relate local pitch movements to a speaker's average voice setting, and it became part of an integral set of notions in the development of a melodic model for the complex phenomenon of sentence intonation.

Given the origin of declination as a useful operational construct, it has since developed into a more ambitious, theoretical concept. Recently, COOPER and SORENSEN [1977, 1981], PIERREHUMBERT [1979], SORENSEN and COOPER [1980] and THORSEN [1980] have made various (psycho-)linguistic claims about how declination correlates with utterance duration, syntactic bracketing and sentence type. These authors claim that certain attributes of declination, such as its rate and resetting, have to be pre-programmed in some detail. This, in turn, commits some authors to the view that declination is under the speaker's conscious control in such a refined way that it can no longer be accounted for by a relatively simple physiological mechanism, i.e. decreasing subglottal air pressure, but must, together with the other pitch movements, be attributed to laryngeal muscle activity.

Apparently, the notion of declination has been allowed to drift from its original moorings. Far from this being a development we have to deplore, it has encouraged us to examine whether the fairly humble status we originally attributed to declination must indeed be elevated.

In the following sections, constituting the main body of our contribution to the debate on the character of declination, we will first give a survey of the original views, then briefly present recent developments and conclusions that have been drawn from these, and finally we will evaluate some of the more ambitious claims regarding its status. For some of these claims we can indeed adduce additional support, but with respect to some others we will show that they lack as yet the necessary experimental evidence.

1. Original Views

1.1. Acoustic Observations

Even before the advent of reliable F_0 meters, the tendency of the voice to lower the medium pitch line during the sentence could be established by means of the unaided ear [PIKE, 1945]. BOLINGER [1964], although probably only referring to the final part of declarative utter-

Fig. 1. F_0 curve of the spontaneously spoken English utterance 'However, help is at hand, perhaps' (female speaker), in which declination lines can fairly easily be drawn.

Fig. 2. Superposition of 11 different utterances, about equally long, by 1 speaker [from MAEDA, 1976]. Dashed lines: fits through the absolute maxima and minima; solid lines: fits through the majority of maxima and minima. The lower solid line had already been drawn by MAEDA [1976].

ances, drew attention to this 'running down pattern' and he hypothesized it to be a language universal.

The introduction of objective F_0 meters does not automatically reveal the declination phenomenon. Figure 1 happens to be a fortunate recording, in which declination is clearly visible, even without the two lines drawn as they have been.

In our own attempts to model F_0 contours of Dutch, we have experienced that the introduction of the concept of declination allowed us to interpret the curves as consisting of two components: one is the declination line, the other is the set of local F_0 rises and falls for which the declination line serves as a reference.

The declination lines drawn in figure 1 are not an accidental property of that particular F_0 curve. The effect of declination is far more systematic, as can be seen in figure 2 [derived from MAEDA, 1976, fig. 2.4.e], which shows the superposition of 11 different utterances of approximately the same length, and spoken by one individual. Evidently, all these curves share about the same downward tendency. To visualize this tendency one can consider drawing one or more possible declination lines through the ensemble. One can fit a line through the absolutely lowest, respectively, highest, F_0 values at the onset and the offset of the utterances. These two dashed lines have the major disadvantage of not passing through the great majority of the local minima, respectively maxima, and are therefore not an optimal representation of a reference line. The solid lines, on the other hand, do connect most of the minima and maxima. The only points missed are the initial dip in the F_0 curve and the exceptionally high first peak.

In some form or other, the representation of an F_0 curve in terms of a superposition of local F_0 rises and falls on a declining line (or between two declining lines) has been fruitful in interpreting the F_0 curves in various languages, e.g. American English [MAEDA, 1976; PIERREHUMBERT, 1981], French [VAISSIÈRE, 1971; DELGUTTE, 1978], Japanese [FUJISAKI and SUDO, 1971], Danish [THORSEN, 1980], Swedish [BRUCE, 1977] and most recently British English [DE PIJPER, 1980]. According to several authors some tone languages exhibit a phenomenon called 'downdrift'. To the extent that this phenomenon is distinct from cases of tonal assimilation, it may be considered comparable to declination [MEYERS, 1974].

Probably a first attempt to establish an empirical rule for the slope of the (lower) declination line was made by MAEDA [1976]. He had

observed that among recordings of F_0 of a given speaker the offset frequency shows only very little variation. He further observed that the differences between onset and offset frequencies are about constant irrespective of the duration of the utterance (which varied between about 1.7 and 2.8 s). Consequently, the rate (r) of declination becomes greater with decreasing utterance length. MAEDA's formula is:

$$r = \overline{\Delta F}/t,$$

where $\overline{\Delta F}$ is the average value for the individual speaker (some 20–30 Hz for male speakers). FUJISAKI and NAGASHIMA [1969] propose two formulae which do not compute the declination rate separately but nevertheless incorporate the declination rate in the specification of the overall contour. COHEN and 'T HART [1967] originally applied the simple recipe of a fixed percentage of fall-off of F_0 per time unit.

1.2. *Physiological Considerations*

The natural tendency of F_0 to gradually fall off in the course of time must originate from some particular property of the mechanism which regulates the rate of vocal fold vibration. That property may be assumed to be particular in the sense that its effect shows up independently of the action of those components of the mechanism that are responsible for the local rises and falls of F_0 .

Compared to normal breathing, speaking requires prolonging the duration of exhalation. This is possible by activity of the external intercostal muscles which serves to prevent the normally rather quick decrease of the size of the thoracic cavity [LADEFOGED, 1967, p. 14]. Due to the expenditure of air, however, the subglottal air pressure P_s decreases rather rapidly, so that phonation becomes no longer possible. As a countermeasure, the speaker diminishes the thoracic cavity by simply relaxing the external intercostal muscles. Apparently, the balance between decreasing P_s and this countermeasure is such that there remains a slow decrease of P_s . Since, other things being equal, F_0 is proportional to the pressure drop over the glottis, and the intra-oral pressure is constant (if averaged over several hundreds of milliseconds), F_0 is proportional to P_s .

The rather scanty physiological data available seem to indicate that this component of F_0 control and the other one, i.e. vocal fold tension, have unique, not overlapping responsibilities: As a first approximation, the gradual decrease of P_s is responsible for the declination pheno-

menon, whereas the laryngeal musculature regulates the variable vocal fold tension which results in momentary rises and falls.

Although it is not clear why, in speech, the balance mentioned above is such that F_0 falls slowly (after all, a singer can easily maintain constant pitch over a considerable duration), this is still a rather simple model. Nevertheless, it accounts for a couple of observations made by COLLIER [1975a]: In two utterances containing a relatively long stretch of baseline declination, the rate of F_0 fall correlated well with the rate of P_s decrease since the $\Delta P_s / \Delta F_0$ ratio corresponded with the generally accepted value of $1/5$ [LADEFOGED, 1967, p. 49]. Laryngeal muscle activity in these utterances did not correlate in any straightforward way with the observed declination.

MAEDA [1976] proposed a variant of this rather simple model. In his data the amount of declination was larger than could be accounted for by the estimated (but not actually measured) decrease of P_s . Therefore, MAEDA assumed an additional, mechanical connection between the decreasing lung volume and declination: The collapsing lungs gradually pull down on the trachea and this tracheal pull causes a rotation of the cricoid cartilage, which results in a progressive shortening of the vocal folds and, hence, in a gradual decrease of their rate of vibration. MAEDA has tested his tracheal pull hypothesis in two utterances and found no reason to reject it.

The hypotheses of both COLLIER [1975a] and MAEDA [1976] agree in attributing the declination phenomenon to a property of the respiratory system. The electromyographic data produced by these authors contain no indication that the laryngeal musculature is in any way involved in the control of declination.

1.3. *Perceptual and Linguistic Relevance*

The perceptual importance of declination becomes evident in any attempt to electronically synthesize F_0 contours: Absence of declination sounds very unnatural (no formal experiments). This may explain why in early synthesis programmes [MATTINGLY, 1966], steps were already taken to simulate the overall downdrift of F_0 in an utterance, even if the concept of declination was not explicitly acknowledged in the approach. In MATTINGLY's programme declination is implemented in a twofold way: (a) F_0 falls over a larger interval than the rise that precedes them, so that the successive peaks and valleys become progressively lower; (b) in those parts of the contour where no rises or falls

occur, F_0 drops very gradually. It should be noticed that provision (b) is perceptually indispensable: When there are no major changes of F_0 over a longer stretch of the utterance, the absence of declination is disturbing and diminishes the impression of naturalness.

It is our basic assumption that in general perceptually relevant pitch movements reflect discrete commands in time on the part of the speaker, and thus may also be relevant in the (psycho-)linguistic domain. Consequently, (certain aspects of) declination too may have linguistic relevance.

2. Recent Developments

2.1. Acoustic Observations

As we have seen in the preceding section, the phenomenon of declination, originally observed as a general tendency of the average fundamental frequency to drop slowly during an utterance, has been given more substance in the sense that declination lines were actually drawn, and formulae for their slope were given. Although figures 1 and 2 suggest that one can observe an upper and a lower declination line in each F_0 recording, it seems to have been the case that the lower one has been held in somewhat more respect than the higher one. MAEDA [1976], for instance, mentions a baseline as the declination line, and is less specific about position and slope of the so-called plateau.

Recently, some authors seem to have felt the need to turn the tables, in favour of the higher declination line. BRECKENRIDGE and LIBERMAN [1977] emphasize the fact that the top line declines at a faster rate than the bottom line, so that the two lines tend to converge. For SORESENSEN and COOPER [1980] this observation is a compelling reason to describe the two lines independently. Moreover, to connect the peaks by means of a top line would be more feasible than to do the same for the valleys in between (although PIERREHUMBERT [1981] also complains that drawing the topline is sometimes not so easy to do). SORESENSEN and COOPER [1980] present a new formula, the Topline Rule, which predicts the height of any peaks intermediate between the first and the last in an utterance.

2.2. Physiological Considerations

BRECKENRIDGE [1977] advances the hypothesis that declination is regulated in the larynx. Her line of reasoning is as follows. The declina-

tion rate is pre-programmed by the speaker, who uses a look-ahead strategy, taking into account utterance length and syntactic phrasing; hence declination is part of the linguistic code. Since all linguistically significant F_0 distinctions are implemented using the laryngeal muscles, declination must be controlled by these muscles, too. This hypothesis is, of course, clearly at variance with the one put forward in section 1.2.

2.3. *Perceptual and Linguistic Relevance*

Two studies seem to support the hypothesis that the declination rate is pre-programmed by the speaker. BRECKENRIDGE and LIBERMAN [1977] found that the second peak in a contour containing two peaks should be produced somewhat lower in order to be heard equally high as the first one. This can be interpreted as an indication of the shared knowledge of speaker and listener of the rate of the topline declination for the (test) utterance at issue. THORSEN's [1980] results point into the direction of a very specific form of pre-programming: In Danish, the declination slope is dependent on sentence category. It is steepest in the case of statements, less steep in continuations or syntactically marked questions and virtually nil in syntactically unmarked questions. Since such linguistic distinctions seem to be encoded in declination, it would be hard to see how an automatic by-product of the respiratory regulation during speech could account for them.

Disregarding the way in which declination is produced, and taking into account declination merely at the acoustical level, SORENSEN and COOPER [1980] derive from the relation between its slope and utterance duration the psychological implication that there is a look-ahead mechanism on the part of the speaker. In their view, different from MAEDA's [1976], not only the slope of the declination, but also its onset value depends on utterance duration: Given a long sentence, a speaker starts at a higher F_0 than if he was about to produce a short sentence. They make the proviso that the look-ahead mechanism does not always work flawlessly in spontaneous speech.

Another implication looms up if, as SORENSEN and COOPER [1980] do, one asks over what domain a speaker applies declination. Considering that the syntactic unit over which the declination extends need not be a full sentence in all cases, one can assume that especially longer sentences are split into two or more parts, each with its own declination. And suitable candidates for such parts would be clauses. At clause boundaries, therefore, 'resetting' of the declination line would

be expected. COOPER and SORENSEN [1977] have the more specific expectation that it may occur at the boundary of conjoined main clauses, but not between main and embedded clauses. They then found no evidence for this hypothesis, but in SORENSEN and COOPER [1980] they attempt a further verification.

They first define 'complete resetting' as the case in which each clause has an identical declination, with the same slopes and starting values of the topline. It would be most evident in a comparison of the F_0 values of the peak prior to the boundary (P_k) with the peak of the word following the boundary (P_{k+1}). The latter should then be significantly higher than the former one, in fact as high as the very first peak in the preceding clause. SORENSEN and COOPER mention that in conjoined main clauses they have indeed observed this indication of topline resetting. They concede, however, that a higher F_0 value for P_{k+1} may also be caused by an extra prominence placed on that word, in the absence of resetting.

Furthermore, 'partial resetting' may also occur: A new declination line begins in the second clause, but with a decreased slope and/or a decreased starting value. In the latter case, too, P_{k+1} is predicted to be higher than P_k , but not as high as the first peak in the first clause.

SORENSEN and COOPER [1980] also propose an improved test of resetting: Several peaks before the main clause boundary should be compared to their counterparts after the boundary and in each matched pair the second peak should be as high as the first. In COOPER and SORENSEN [1981] a variant of such an improved test is performed, with two matched pairs of sentences (Experiment 2.3.2, pp. 91 ff.). In one sentence, four successive peaks belong to the same syntactic constituent; in the other one the first peak occurs before a main clause boundary, whereas the next three are at the beginning of the second conjunct. According to the hypothesis, the latter three peaks should be systematically higher than their counterparts. The prediction is borne out for the first two peaks following the boundary, but not for the third. Therefore, COOPER and SORENSEN reject the strong version of their hypothesis.

3. *Present Outlook: Critical Evaluation*

3.1. *Acoustic Observations*

On several grounds, it is preferable to plot F_0 recordings using a

logarithmic frequency scale. For one thing, a logarithmic scale is in better agreement with the way in which pitch differences are perceived than a linear scale. For another, as a consequence of this, changes of F_0 which sound similar, also show up similar, irrespective of the average overall height. This enables us to compare F_0 recordings of speakers with largely different average voice height, such as males and females.

If a logarithmic scale is applied, the convergence of top and bottom lines decreases to almost nil in most cases (as in fig. 1). Admittedly, even with a logarithmic scale the convergence remains visible in a number of cases, but auditorily there is no difference with two parallel lines.

In our opinion, therefore, there is no need to consider the slopes of the top and bottom lines to be independent. The bottom line can be seen as a lower parallel to the top line, or the top line as a higher parallel to the bottom line. We have preferred the latter view, and still do so. One reason is that if one draws two declination lines in an F_0 curve and if one repeats this operation for a number of utterances (of about equal length and produced by the same speaker), and next superimposes these sets of lines, it turns out that the bottom lines overlap considerably, whereas the top lines appear at varying heights above the baseline. In other words, the bottom line is a more stable reference line than the top line, which varies as a function of the intervals covered by the F_0 rises (and falls). So, we would propose a certain subordination of the two lines, such that the bottom line has a relatively fixed location within the speaker's pitch range, whereas the top line runs above it in parallel and at a variable distance, controlled by the independent factor of pitch range.

Serious attempts to quantify the slope of the declination line should be based mainly on F_0 recordings in which declination shows up in its pristine form, i.e. in utterances that are not overcrowded with pitch accents. In such recordings, the bottom line is not merely a virtual line which connects the majority of the valleys in the curve, but it is directly visible (fig. 3). In this way, for Dutch the following empirical rule was proposed ['T HART, 1979; 'T HART et al., 1982]:

$$\begin{array}{ll} D_1 = -8.5/t & \text{for } t > 5 \text{ s,} \\ D_2 = -11/(t+1.5) & \text{for } t \leq 5 \text{ s,} \end{array}$$

with D in semitones per second and t in seconds. D_1 predicts a fixed interval between onset and offset frequency and hence a slope which

Fig. 3. F_0 curve of the English utterance 'It's not surprising they never make any money' with only one pitch accent, so that the (lower) declination line following it can actually be seen.

varies with duration. This is roughly the same formula as MAEDA's [1976], except that this formula predicts a less steep slope for any given duration (for male voices). For durations shorter than 5 s D_2 indicates that, with decreasing utterance length, not only the slope is steeper, but also the onset frequency is lower. Exactly the same formulae can be used to account for the declination slopes in British English utterances [DE PIJPER, 1980]. Recently, WILLEMS [1982] has checked the formulae against 35 British English utterances, read aloud by 5 different speakers, male and female, and against another 35, this time spontaneous utterances by 10 different speakers. Durations varied between 0.6 and 6.3 s. It turned out that in both cases the mean slope, averaged over all the utterances, was fairly accurately predicted (deviation -0.51 semitone/s for utterances read aloud, -0.30 for spontaneous ones), but that the moderate standard deviation in the loud reading condition (of 0.79 semitone/s) was almost tripled in the spontaneous material. Figure 4 shows the distributions of the deviations, solid for utterances read aloud, dashed for spontaneous utterances.

Both D_1 and D_2 suggest that declination proceeds linearly (i.e. $\log F_0$ as a function of time is a straight line). We feel justified in saying

Fig. 4. Frequency distributions of the differences (in semitones/s) between declination slopes calculated on the basis of Dutch material and actually occurring in British English material. Solid lines: 35 sentences read aloud; dashed lines: 35 spontaneous utterances. The means are fairly close to the predicted values: -0.51 semitone/s for loud reading, -0.30 for spontaneous material; standard deviations are 0.79 and 2.05 semitones/s, respectively. Durations varied between 0.6 and 6.3 s.

that this is a sufficiently accurate approximation for short utterances. However, it is unknown whether such a linear approximation is also realistic for longer utterances. Two alternatives, apart from resetting, seem at least possible: Either a rather steep slope in the beginning and a levelling off towards the end, or little or no declination in the beginning and starting the genuine declination only when the end of the utterance can be anticipated (which would be a rather trivial look-ahead strategy).

Fig. 5. Schematized representation of two contours of which A sounds terminal and B does not, despite A having a smaller final fall. In A, the fall lowers F_0 to the baseline, which is not the case in B.

These reservations about the actual implementation of declination in longer utterances, especially in spontaneous speech, have to be made since there is no sufficient evidence so far that a look-ahead principle systematically controls the slope of the declination line in anything but short utterances read aloud. Although the material examined by SORENSEN and COOPER [1980] suggests that their Toplevel Rule is satisfied fairly accurately, their claim about a look-ahead mechanism is not supported by sufficiently varied observations [SIMON, 1980, pp. 542–543].

3.2. *Physiological Considerations*

BRECKENRIDGE's [1977] line of reasoning in favour of laryngeal control of declination is debatable, since there is not enough evidence to say that all linguistically significant F_0 distinctions are implemented using the laryngeal muscles. Since, moreover, we cannot be very sure about pre-programming and look-ahead strategies other than in short sentences read aloud, it is an undue generalization to declare that declination is part of the linguistic code.

Nevertheless, in view of the possibility of resetting and of reduction or absence of the declination, we might ask how these deviations of the usual behaviour of declination could be implemented by other means than the laryngeal muscles. Two alternative explanations offer themselves. One is that the speaker directly reduces or suppresses declination by the appropriate regulation of his P_1 . The other is that the speaker

regulates his P_s in invariably the same way, but overrides the declination effect with the appropriate amount of laryngeal muscle tension. There is some indication that the latter account is more plausible: In combined EMG measurements in utterances with 'no declination' or 'inclination' increased cricothyroid muscle activity correlates far better with the F_0 variation than does the fairly constant P_s [COLLIER, 1975a].

3.3. *Perceptual and (Psycho-)Linguistic Considerations*

The following can be looked upon as more evidence for considering the bottom line as the more basic reference. In a number of languages a distinction can be made between two intonational categories (patterns), one of which is associated with 'statement' and the other with 'continuation' or 'calling' (under certain conditions). Inspection of the corresponding F_0 curves shows that both end in a fall, but gives the impression that the fall is larger in the statement pattern. An appropriate difference in size is indeed sufficient to elicit the corresponding categorical distinction in listeners [COLLIER, 1975b, pp. 298ff.).

So, a first explanation of the effect could be based on either of the following two hypotheses: (a) the absolute size of the fall is the decisive cue; (b) not the absolute size of the fall, but its size in relation to the preceding rise is the decisive cue, in the sense that the statement pattern requires a fall that covers approximately the same interval as the rise.

Figure 5 schematically represents an observation which refutes both hypotheses. Contour A is interpreted as a statement, despite its having a fall that is smaller than the preceding rise. Contour B sounds as the other pattern, even if its fall is considerably larger than the preceding rise.

An alternative explanation can be based on the consideration that the fall in A lowers F_0 to the level of the lower declination line, whereas the fall in B fails to do so. In other words, the listener evaluates the category of the fall with reference to his mental projection of the lower declination line. *Mutatis mutandis* the same must hold for the speaker.

The mental projection of the lower declination line allows for a full explanation of the outcome of the experiment by BRECKENRIDGE and LIBERMAN [1977], as reported in PIERREHUMBERT [1981]: The second peak in their stimulus was elevated as much above the – mentally projected – lower declination line as was the first one, and for that reason it is judged to be equally high, although the peak frequency is lower than that of the first peak.

From the test of our declination slope formulae we may conclude that when reading aloud isolated sentences the speaker can literally look over the text and pre-programme the declination slope as a function of utterance length. The same might hold for spontaneously produced utterances that are short enough for the look-ahead mechanism to cope with, but anyway the formulae do not work very accurately for individual spontaneous utterances, even if they are short.

Formula D_1 is intended for longer utterances of more than 5 s. It should be observed that this formula is already less specific than D_2 . It does not include a correlation between onset F_0 value and utterance length. It only states that the overall decrease of F_0 due to declination is 8.5 semitones.

As we have said in the 'Introduction', studying acoustic recordings of F_0 is not sufficient to allow us to lay bare the way in which variations of F_0 function in the melodic structure of the intonational system of a language. Apart from the general need of perceptual evaluation of supposed effects in the acoustic measurements, this necessity seems to exist particularly with respect to the potentially interesting issue of resetting. COOPER and SORENSEN [1981] have measured statistically significant peak height differences restricted to within 6 and 13 Hz. We doubt whether this is a large enough cue for the listener.

In their 1977 paper (their fig. 1), COOPER and SORENSEN produced an F_0 curve in which there is no resetting of the topline declination, because the F_0 peaks on either side of the syntactic boundary are of roughly equal height. An alternative way of interpreting this curve is to concentrate on bottom line declination: This can be seen to contain an F_0 difference of at least 30 Hz, which we would guess to be perceptually relevant, so that it could be taken as a strong indication of bottom line resetting.

Apart from the claims on declination resetting, we have met yet another linguistically inspired hypothesis, i.e. THORSEN's [1980] dependence of the declination slope on the syntactic category of the sentence. The fact that this hypothesis has been confirmed, also in perceptual experiments, is surprising, since in formulas for declination slope put forward by others, utterance duration is the only relevant parameter. One may well ask whether Danish is the only language in which declination slope depends on sentence category but not on utterance duration, or else, whether it will be found to be an additional parameter in other languages.

4. Conclusion

In this final section we will recapitulate our position with regard to the various issues dealt with above, and we will present some suggestions for further research. Some authors claim that two declination lines have to be distinguished (a topline and a bottom line), the main argument being that the two lines may have different slopes. We have argued that one declination line is sufficient because in the logarithmic domain the top and bottom line tend to run parallel, so that one can be derived from the other.

If one of the two lines has to be chosen as the basic reference line, the bottom line is the preferred candidate. Acoustic arguments in favour of this decision are: (1) superimposed F_0 traces of declarative utterances by one speaker show a substantial overlap (and hence invariance) of the bottom lines; (2) when the curve contains only one peak, no topline can be drawn at all, whereas the bottom line then shows up in its pristine form. Perceptual arguments are: (1) only the bottom line can account for the interpretation of 'finality' as cued by pitch falls irrespective of their interval; (2) bottom line declination is a more elegant explanation of the way in which listeners evaluate height differences of successive peaks (thus providing an alternative to PIERREHUMBERT's [1981] account of that mechanism).

Our formulae for the slope of declination require less input data and have more predictive power than SORENSEN and COOPER's [1980], Topline Rule: To calculate the slope we only need to know what the utterance duration is, and the formulae are applicable irrespective of the number of peaks (the Topline Rule does not work for contours with only one single peak). In order to actually draw the baseline, the only additional information needed is the terminal F_0 value – which is fairly invariable for an individual speaker.

As far as the physiological explanation of declination is concerned, we have proposed a relatively simple model in which properties of the respiratory mechanism are seen as the main cause. Admittedly, there is as yet no conclusive evidence in favour of this model, but even if this has not yet been found we nevertheless think that there is no good reason to abandon it and adopt a more specific model that puts declination under laryngeal control.

Discussing the issue of pre-programming of the declination slope we have given evidence that the postulated look-ahead mechanism ap-

parently operates under rather restricted conditions only, i.e. if short utterances are read aloud. Accordingly, we have proposed declination formulae that make less specific predictions as utterance length increases. Moreover, for spontaneous speech our formulae do not accurately predict the slope of individual utterances, but are indicative only of the slope averaged over a number of utterances.

As far as possible correlations between declination resetting and syntactic properties of utterances are concerned, this brings us to suggestions for further research, since this potentially interesting point – in that it may have particular communicative relevance – has only begun to attract the attention of some authors.

Instead of starting from the primacy of linguistic distinctivity, it seems wise to start the other way round, i.e. to search in F_0 curves for (base line) resetting (and other ways of splitting up the declination line) in such a way that first the question is answered whether it happens, and subsequently where it happens.

We do not believe that further attempts to refine the proposed formulae for declination slope will be very fruitful. Quite another point, however, is whether the large individual deviations from the predictions as have been reported are haphazard or if, also in other languages than Danish, there is a systematic correlation with some non-physiological determining factor. About the question of which shape the declination might have if a linear approximation in longer utterances without resetting is not adequate, we can say that our own experience with a rather high rate in the beginning and a gradual levelling off towards the end have been disappointing from a perceptual point of view, so that we decided to opt for a straight line (in the logarithmic F_0 domain). This does not, of course, exclude the possibility that an initial moderate rate and a steeper final part is still a better solution.

It seems desirable to do combined measurements of F_0 , P_s , and EMG of selected intercostal muscles to learn more about the possibly exclusive role of respiratory activity in declination. Rather than attempts to find new evidence for look-ahead strategies, two types of experiments seem promising. One is about the question of whether listeners' interpretations can be influenced by manipulation of the location of resetting. The other one could concentrate on the question of whether the mental projection of the baseline can be interfered with when different declination slopes from the ones expected by the listener are applied to the stimulus material.

In summary, the notion of declination, originally introduced as a useful tool to interpret F_0 curves more easily, has recently been examined with regard to its potentially more important status as a theoretical concept in intonation theory. We have welcomed this new interest in an old phenomenon as an occasion to redefine our own position and to evaluate some of these recent proposals. In the process we have indeed found reasons to attach greater psycholinguistic importance to declination than we used to do. At the same time we have pointed out that some recent claims, though interesting in themselves, lack as yet conclusive evidence.

References

- BOLINGER, D. L.: Intonation as a universal. Proc. IXth Int. Congr. Linguists, Cambridge 1962 (Mouton, the Hague 1964).
- BRECKENRIDGE, J.: Declination as a phonological process. Bell Laboratories Technical Memorandum (Murray Hill 1977).
- BRECKENRIDGE, J.; LIBERMAN, M. Y.: The declination effect in perception. Bell Laboratories Technical Memorandum (Murray Hill 1977).
- BRUCE, G.: Swedish word accents in sentence perspective (Gleerup, Lund 1977).
- COHEN, A.; HART, J. 'T: Perceptual analysis of intonation patterns. Proc. Vth ICA, Liège 1965, A. 16.
- COHEN, A.; HART, J. 'T: On the anatomy of intonation, *Lingua* 19: 177-192 (1967).
- COLLIER, R.: Physiological correlates of intonation patterns. *J. acoust. Soc. Am.* 58: 249-255 (1975a).
- COLLIER, R.: Perceptual and linguistic tolerance in intonation. *IRAL XIII*: 293-308 (1975b).
- COOPER, W. E.; SORENSEN, J. M.: Fundamental frequency contours at syntactic boundaries. *J. acoust. Soc. Am.* 62: 683-692 (1977).
- COOPER, W. E.; SORENSEN, J. M.: Fundamental frequency in sentence production (Springer, New York 1981).
- DELGUTTE, B.: Technique for the perceptual investigation of F_0 contours with application to French. *J. acoust. Soc. Am.* 64: 1319-1332 (1978).
- FUJISAKI, H.; NAGASHIMA, S.: A model for synthesis of pitch contours of connected speech. Annu. Rep., Eng. Res. Inst., Fac. Eng., Univ. Tokyo, No. 28, pp. 53-60 (1969).
- FUJISAKI, H.; SUDO, H.: Synthesis by rule of prosodic features of connected Japanese. Proc. VIIth ICA, vol. 3, pp. 133-136 (Akadémiai Kiadó, Budapest 1971).
- HART, J. 'T: Explorations in automatic stylization of F_0 curves. IPO Annu. Progr. Rep., No. 14, pp. 61-65 (1979).
- HART, J. 'T; NOOTEBOOM, S. G.; VOGTEN, L. L. M.; WILLEMS, L. F.: Manipulations with speech sounds. *Philips tech. Rev.* 40: 134-145 (1982).
- LADEFOGED, P.: Three areas of experimental phonetics (Oxford University Press, Oxford 1967).
- MAEDA, S.: A characterization of American English intonation; PhD thesis MIT Cambridge (1976).
- MATTINGLY, I. G.: Synthesis by rule of prosodic features. *Lang. Speech* 9: 1-13 (1966).
- MEYERS, L.: Tone patterns in Hausa: a re-analysis of Hausa downdrift. UCLA Working Papers in Phonet., No. 27, pp. 47-62 (1974).

- PIERREHUMBERT, J.: The perception of fundamental frequency declination. *J. acoust. Soc. Am.* 66: 363-369 (1979).
- PIERREHUMBERT, J.: Synthesizing intonation. *J. acoust. Soc. Am.* 70: 985-995 (1981).
- PIKE, K. L.: The intonation of American English (University of Michigan, Ann Arbor 1945).
- PIJPER, J. R. DE: A melodic model of British English intonation. *IPO Annu. Progr. Rep.*, No. 15, pp. 54-58 (1980).
- SIMON, H. A.: How to win at twenty questions with nature; in COLE Perception and production of fluent speech, pp. 535-548 (Lawrence Erlbaum Ass., Hillsdale 1980).
- SORENSEN, J. M.; COOPER, W. E.: Syntactic coding of fundamental frequency in speech production; in COLE Perception and production of fluent speech, pp. 399-440 (Lawrence Erlbaum Ass., Hillsdale 1980).
- THORSEN, N.: A study of the perception of sentence intonation - evidence from Danish. *J. acoust. Soc. Am.* 67: 1014-1030 (1980).
- VAISSIÈRE, J.: Contribution à la synthèse par règle du Français; thèse Université de Grenoble (1971).
- WILLEMS, N.J.: Towards an objective course in English intonation: standardized precepts; in Studies for Antonie Cohen (Foris, Dordrecht, in press 1982).

Received: July 7, 1982

Prof. A. COHEN, Institute of Perception Research (IPO),
PO Box 513, NL-5600 MB Eindhoven (The Netherlands)