Working Papers 21 - 1981
Linguistics - Phonetics
Lund University

TONAL AND TEMPORAL INTERPLAY

Gösta Bruce

INTRODUCTION

When in my thesis I discussed the interplay of tonal commands, I assumed a basically fixed temporal program (Bruce 1977, chapter 5). The observed tonal adjustments, e.g. interruption, anticipation and delay of Fo-changes, were interpreted as the result of an adaptation to the given time restrictions in cases where tonal commands conflict. This tonal dependence on the temporal setting is presupposed in the intonation models in Bruce (1977, chapter 8) and Bruce and Gårding (1978). The rationale for this view is that certain rhythmical demands connected to the concept of isochrony must be fulfilled.

The same view that the temporal program is basic and the tonal program has to adjust to it finds support in a study by Erikson (1973) on the synchronization of the onset of the secondary stress syllable and the Fo-contour in Swedish disyllabic compounds. An increase in the number of consonants in the medial cluster and thus a delay of the second vowel onset will cause a corresponding delay of the Fo-contour of this syllable. In their modeling of the temporal organization of Swedish utterances Lindblom et al. (1976) also imply a temporal independence from tonal events.

Recent work on tonal and temporal interplay lends support to a different view of this dependence. In a study of final lengthening, Lyberg (1979) suggests that the final lengthening observed in Swedish may at least partly be a function of the fundamental frequency change required in final position. This is to say that tonal demands might govern the durations of the segments involved.

Öhman et al. (1979) hypothesize that the actual durations of segments may be secondary effects of other, "primary acoustic effects". For example they have found that an Fo-change covering a wide range tends to lengthen the segments involved.

In their extreme form these two hypotheses about tonal and temporal interplay make opposite predictions (cf. Lindblom's discussion in Nordic Prosody 1978, pp. 293-295). Figure 1 is an attempt to illustrate the consequences of the hypotheses for the tonal and temporal domains. For each

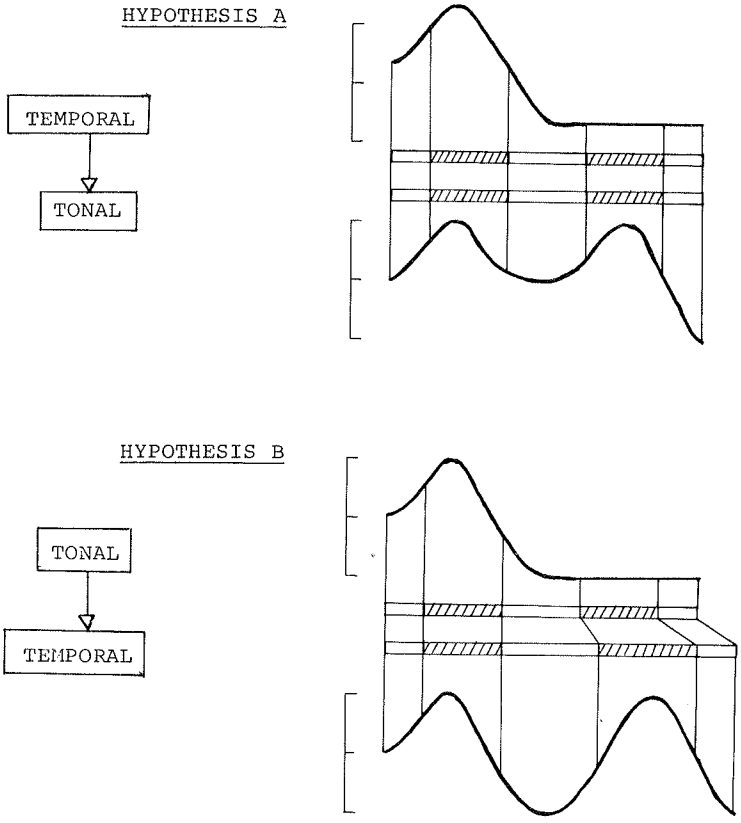

Figure 1. Illustration of consequences of two opposite hypotheses about tonal and temporal interplay for the Fo-contour and segment durations. For each hypothesis a simpler (upper part) and a more complex (lower part) Fo-contour is shown.

hypothesis a simpler (upper part) and a more complex Fo-contour are shown. These hypothetical Fo-contours are similar to those found in a Standard Swedish accent 2-word in non-focal and focal position respectively.

According to hypothesis A (temporal-to-tonal) the addition of a second rise-fall after the first one forces Fo to make a shortcut - interruption of the Fo-fall by the immediately following rise - to fit a fixed temporal program. Hypothesis B (tonal-to-temporal), on the other hand, predicts that adding another rise-fall to the first one will cause a lenghtening of the

actual segments.

The purpose of the present study is to study the tonal and temporal interplay in the light of these two extreme hypotheses. It should be pointed out from the beginning that they serve as null hypotheses against which this interplay can be studied. One way of doing this, the one that will be used here, is to keep the segmental context constant (or nearly constant) and vary the tonal context. Standard Swedish with its relatively complex tonal system offers good possibilities for studying this interplay.

TEST MATERIAL AND INFORMANTS

The test material for this study consists of a subset of the test material used in my thesis (Bruce 1977) and recorded by my main informant. In this material placement of sentence accent was varied (three different focus locations), and for the (first and) final focus location choice of word accent. The following sonorant test sentence are included. The portions within parentheses have not been analysed for this study:

(Nian vill $\begin{Bmatrix} 1 \\ \text{a}\acute{}\text{namma} \\ \text{}\grave{}\text{lämna} \end{Bmatrix}$ nåra) $\grave{}$långa $\begin{Bmatrix} 3 \\ \acute{}\text{nummer} \\ \grave{}\text{nunnor} \end{Bmatrix}$.(One wants to $\begin{Bmatrix} 1 \\ \text{accept} \\ \text{leave} \end{Bmatrix}$ some) $\begin{Bmatrix} 2 \\ \text{long} \\ \text{tall} \end{Bmatrix}$ $\begin{Bmatrix} 3 \\ \text{numbers} \\ \text{nuns} \end{Bmatrix}$.

A similar material was recorded by a second informant:

(Dom vill $\grave{}$slippa) $\grave{}$långa $\begin{Bmatrix} 3 \\ \acute{}\text{nummer} \\ \grave{}\text{nunnor} \end{Bmatrix}$. (They want to avoid) $\begin{Bmatrix} 2 \\ \text{long} \\ \text{tall} \end{Bmatrix}$ $\begin{Bmatrix} 3 \\ \text{numbers} \\ \text{nuns} \end{Bmatrix}$.

In addition a parallel material in an obstruent context, where segmentation difficulties are judged to be less than in a sonorant environment, was recorded by the same informant. Besides varying the placement of sentence accent and choice of word accent in final position in this material, the phrase structure of the last part of the sentence containing either one word (compound) or two words was varied. Variation of the placement of secondary stress in the compound was also included:

Dom vill $\grave{}$slippa $\grave{}$nacka $\begin{Bmatrix} 3 \\ \acute{}\text{skotten} \\ \grave{}\text{skotten} \end{Bmatrix}$. They want to avoid chopping the $\begin{Bmatrix} 3 \\ \text{shoots} \\ \text{Scotsman} \end{Bmatrix}$.

Dom vill $\grave{}$slippa $\grave{}$nacka, skotten. They want to avoid the Nacka-Scotsman.

(Dom vill $\grave{}$slippa) $\begin{Bmatrix} 2 \\ \grave{}\text{nacka, schacket} \\ \grave{}\text{nackajac, kett} \end{Bmatrix}$.(They want to avoid) the $\begin{Bmatrix} 2 \\ \text{Nacka-chess} \\ \text{Nacka-coat} \end{Bmatrix}$.

Both informants are female and represent the Stockholm variety of Standard Swedish.

RESULTS

Figures 2 and 3 show average values of segment durations and Fo-peaks and valleys of sonorant test phrases for the two speakers.

In non-final position the two speakers seem to use different strategies. For the word ˋlånga speaker EH appears to have approximately the same durations in focal (Fig 2, medial part) and pre-focal position (Fig 2, upper part). The higher tonal demands of the focus - because of the sentence accent rise, which is absent out of focus (cf. Bruce 1977, chapter 3) - result in an interruption of the word accent fall by the sentence accent rise in ˋlånga. This is evident for ˋlånga in the context of the accent 2-word ˋnunnor (Fig 2, right column), while the situation in the context of the accent 1-word ˊnummer (Fig 2, left column) is less clear-cut.

Speaker UN, on the other hand, lengthens all segments in ˋlånga in focal position (Fig 3, medial part), as compared to pre-focal position (Fig 3, upper part). This lengthening is non-uniform, however, and affects mainly the duration of the long consonant [ŋ:] in ˋlånga. No interruption of the word accent fall is apparent. In fact, the range of the Fo-fall is wider in focal than in pre-focal position. In particular the extra increase of the long consonant might be explained by reference to the higher tonal demands of the focus,with the sentence accent rise involving mainly this segment.

In final position, however, the speakers use similar strategies. For the accent 2-word ˋnunnor the durations are longer in focal (Figs. 2 and 3, upper part) than in post-focal position (Figs. 2 and 3, medial part). The lengthening is found mainly for the long consonant [n:] and the final vowel [ɔ], exactly those segments where a temporal adaptation to tonal demands should be expected in focal position, as the sentence accent rise and the terminal jucture fall - which are absent in post-focal position - are manifested here (cf. Bruce 1977, chapter 3). The range of the word accent fall is clearly wider in post-focal position, however. In focus the word accent fall in ˋnunnor is interrupted by the sentence accent rise. This means that for both speakers we have a tonal as well as a temporal adaptation in focal position. A parallel picture is found also for the

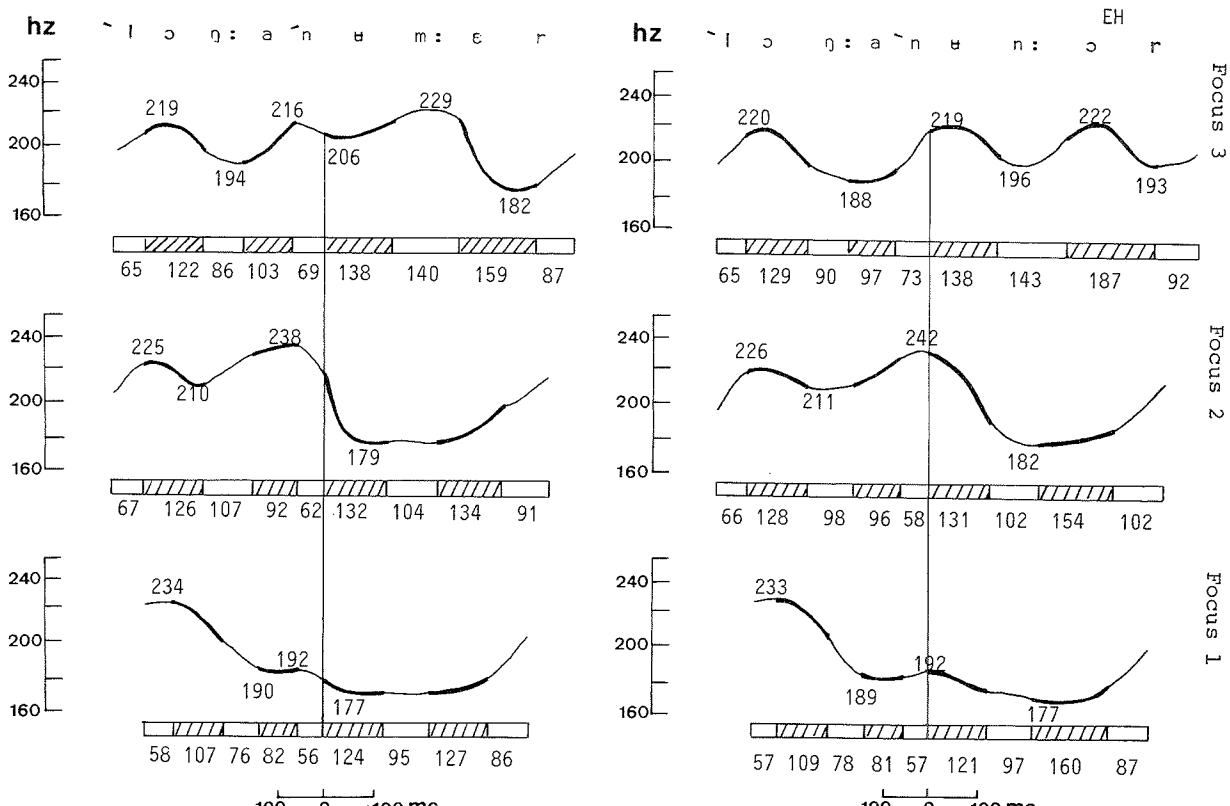

Figure 2. The effect of the placement of sentence accent and choice of word accent. Speaker EH. Typical Fo-contours for two-word phrases in sentence final position. Mean values of 10-14 repetitions of Fo-peaks and valleys and segment durations. The line-up point is at the CV-boundary of the final stressed syllable.

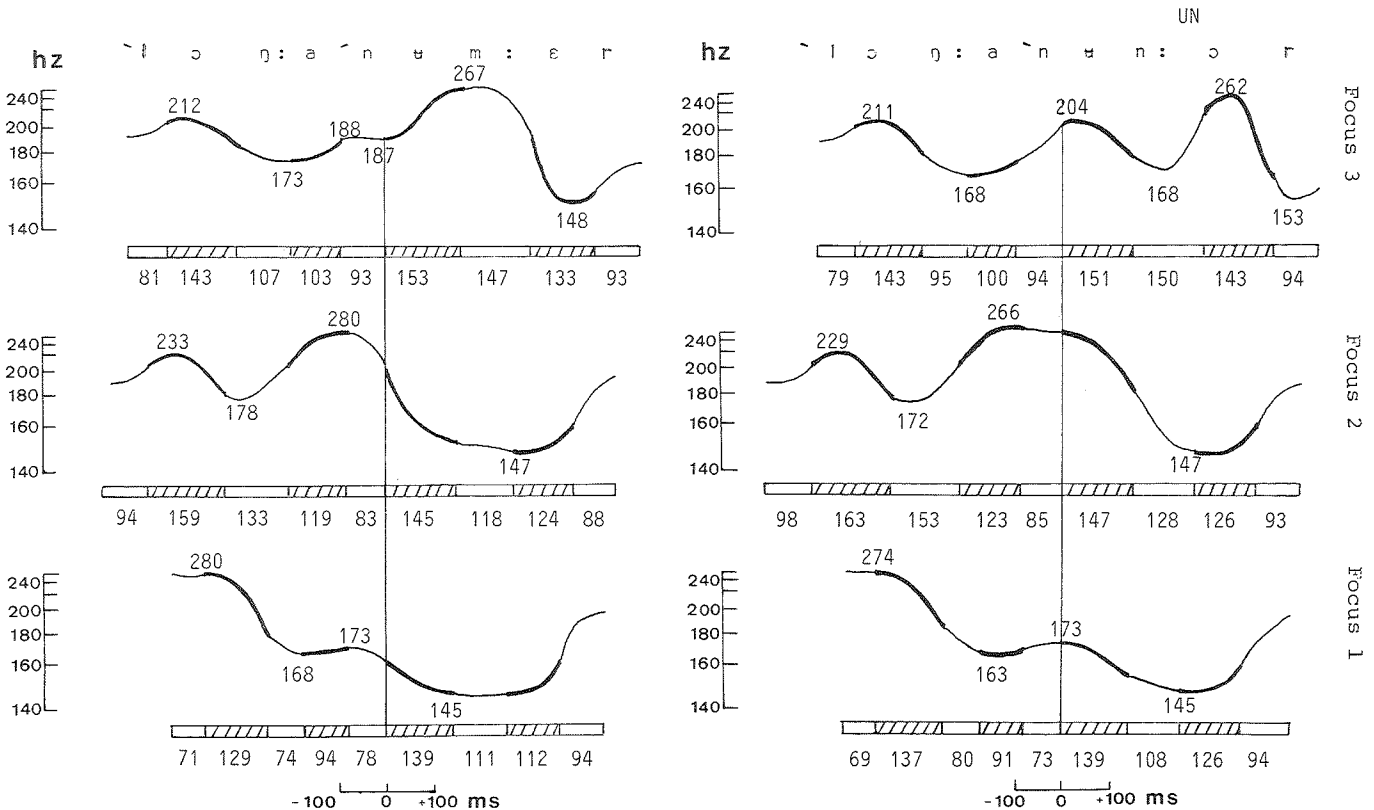

Figure 3. The effect of the placement of sentence accent and choice of word accent. Speaker UN. Typical Fo-contours for two-word phrases in sentence final position. Mean values of 6 repetitions of Fo-peaks and valleys and segment durations.

accent 1-word ´nummer in focal and post-focal position.

According to hypothesis B (tonal-to-temporal) we would expect more lenghtening for an accent 2-word than for an accent 1-word in focus because of the more complex tonal pattern (cf. Figs. 2 and 3, upper part). This prediction is not supported by the data. We find approximately the same durations for accent 1 and accent 2. There are only marginal differences in duration. We would have expected a larger increase of the long consonant and also of the final vowel for the accent 2-word, as compared to the accent 1-word, with reference to the difference in tonal complexity there. These results contradict the findings by Öhman et al. (1979, p. 309) for a similar comparison.

For speaker EH there is a durational difference in the final vowel, which might be tied to the higher tonal demands in the accent 2-word. But comparing the post-focal version of ´nummer/`nunnor (cf. Figs. 2 and 3, medial part), where the Fo-difference for the final vowel is absent, approximately the same difference in duration for the final vowel is present. Therefore it is likely that the vowel difference [ɛ] versus [ɔ] is responsible for the longer duration of the final vowel in the accent 2-word.

The conclusion is then that the same temporal program is used in focus (for both speakers) for accent 1 and accent 2 in spite of the higher tonal demands for accent 2.

That a wider range of an Fo-change does not necessarily bring about longer durations is clearly evident in Figs. 2 and 3, showing the durations of `långa in post-focal (lower part) compared with pre-focal position (upper part). The much wider range of the Fo-fall (for both EH and UN) in post-focal than in pre-focal position is not accompanied by longer durations of the segments involved. In fact the durations are shorter in post-focal position.

For the two post-focal versions of ´nummer and `nunnor (Figs. 2 and 3, medial and lower part) there is a considerable difference in the Fo-range of the word accent fall. In spite of this clear Fo-difference, the increase in the segment durations of the version with a wider range is only marginal.

A case where tonal demands cannot explain temporal differences is exemplified in Figure 4. The temporal structure of the last portion /ʃaket/ of the compound is clearly different, if the secondary stress is on the penultimate or the ultimate syllable. In focal position (Fig 4, left

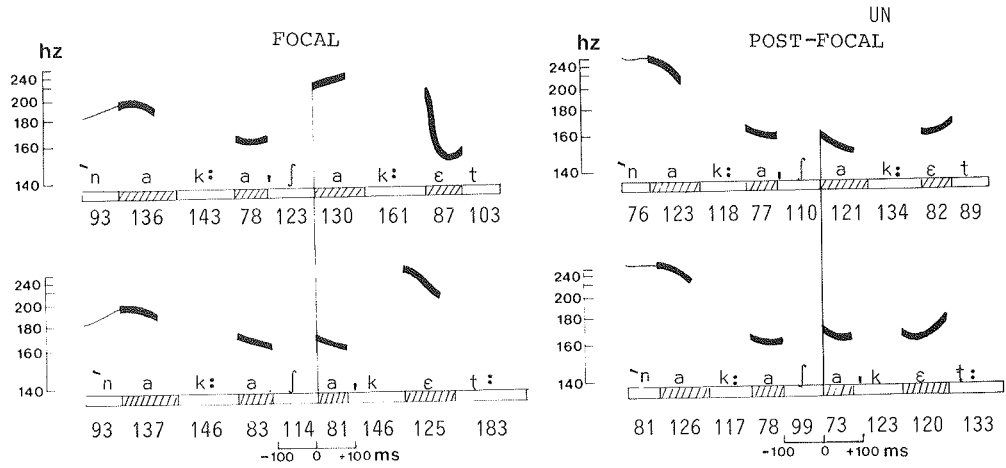

Figure 4. The effect of the placement of sentence accent. Speaker UN. Typical Fo-contours of compounds with varying placement of secondary stress. Mean values of 6 repetitions of segment durations.

column) the Fo-contours are distinct, but in post-focal position (Fig 4, right column) the Fo-contours are more or less identical. In this latter context the secondary stress placement appears to be signalled primarily by the difference in temporal structure (cf. Bruce 1977, p.14).

DISCUSSION

Erikson's paper (1973) on the synchronization of articulatory and phonatory processes cited as evidence for the temporal-to-tonal hypothesis showed a delay in the Fo-contour of the second syllable when the number of intervocalic consonants increases, and consequently the vowel onset of the second syllable is delayed as in words like ˋhingst,språng compared to ˋbi,hang. It should be noted that it is often the case - at least in running speech - that a heavy consonant cluster will be reduced, probably decreasing the duration of the cluster, e.g. [hiŋ:stspr ɔŋ:] pronounced as [hiŋ:spr ɔŋ:]. Although it is fully possible to retain the whole cluster even in running speech, the dropping of consonants is customary in heavy clusters, which may partly be due to the rhythmical constraints of a relative isochrony.

It is clear from this study, however, that the Fo-contour does not live its own life. The Fo-contour - or certain critical Fo-points - are crucially dependent on certain articulatory events. This point is probably agreed on even by those who advocate the tonal-to-temporal hypothesis (cf. Öhman et al. 1979, p. 310) and is not under debate. In this sense the temporal-to-tonal hypothesis is supported.

But it is still not clear , whether the realization of a complex tonal pattern can force Fo to make adaptations to the temporal program, or whether it will elicit a stretching of the segments involved, or some combination of both.

Comparing focal and non-focal position we find a more complex tonal pattern in focus. Normally we also find longer segment durations in focus. This is evident from Figure 5, which gives the durations for speaker UN for the phrase /slip:a nak:a skot:ɛn/. The Figure shows that all segments are being stretched in focus position. The increase is not confined to the segments - mainly the long consonant - where the sentence accent rise is found, although the increase appears to be larger here. This general increase may be interpreted as a direct contribution of focus to the temporal domain. i.e. perhaps both tonal and temporal changes are independent consequences of focus. Temporally this means that the focal part - the most informative part of an utterance - is simply given more time space.

As was noted above, the largest increase in duration for focus position is given to the long consonant of the stressed syllable. It has been pointed out that this is the segment where the sentence accent rise is usually executed, which lends support to the tonal-to-temporal hypothesis. But there is another possible explanation. In a study of the effect of sentence accent on quantity, Bannert (1979) has shown that it is the phonologically long segment of a stressed syllable that is given the largest increase in duration, i.e. the long vowel in a V:C-structure and the long consonant in a VC:-structure. This non-uniform increase in duration - the long segment increasing more than other segments - is expected for the following reason. If the quantity contrast is to be preserved, the durational differences must increase with increasing segment durations (cf. Bannert 1979, p. 256). Therefore the non-uniform increase of segment durations in focus can be accounted for without any reference to specific tonal demands. Moreover the tonal-to-temporal hypothesis cannot be used to explain the

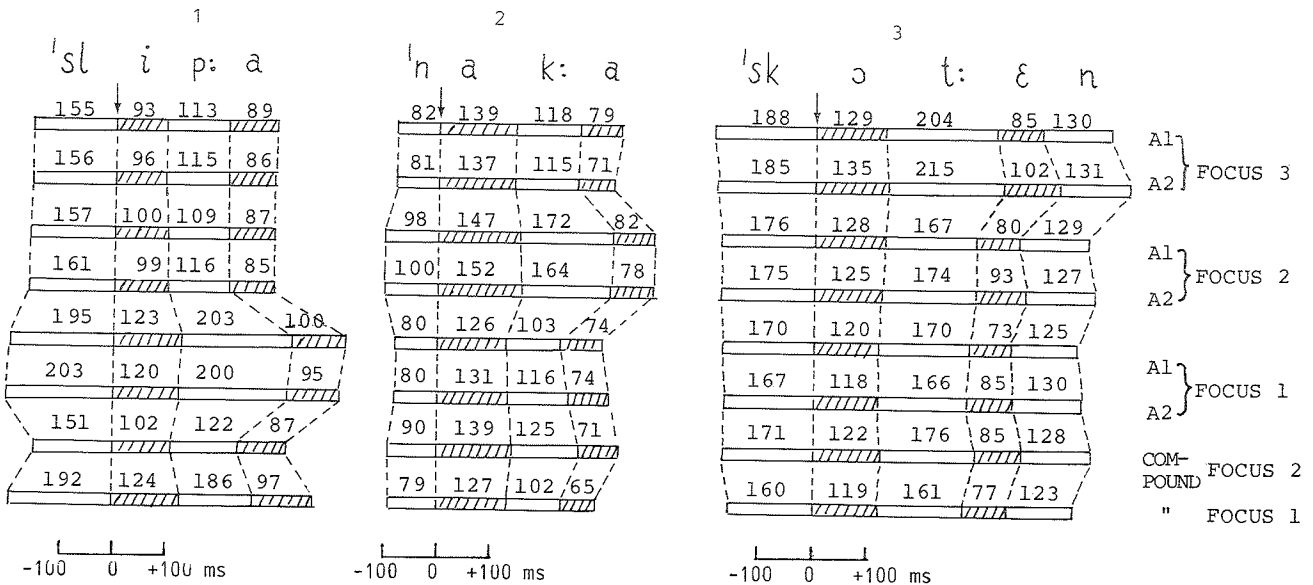

Figure 5. The effect of the placement of sentence accent, choice of word accent and phrase structure of the last part of the sentence (two-word phrase/compound). Speaker UN. Mean values of 10-13 repetitions of segment durations. Line-up points are at the CV-boundary of each stressed syllable.

extra increase of the long segment in a V:C-context, as the long vowel does not directly involve the tonal contribution of focus.

In a situation where focus and position in the sentence are kept constant and only word accent is varied, giving a more complex tonal pattern for the accent 2-word than for the accent 1-word, the temporal differences were shown to be minimal. If hypothesis B (tonal-to-temporal) were true, we would have expected a clear temporal difference between the two tonal contexts. With reference to the above discussion it can be concluded that the evidence for the tonal-to-temporal hypothesis in the present study is scarce.

It is of course possible that the actual temporal program for the focal portion of an utterance in Swedish "has been sculptured by the evolutionary forces of language use" (cf. Lindblom et al. 1976, p. 65). As there is usually a more complex tonal pattern in focal position, this may have contributed to longer durations for the focal portion than for non-focal portions of an utterance. This would mean that we are faced with a temporal preprogramming and not a moment-by-moment adjustment of the temporal program to the specific tonal demands, as has been suggested by Lyberg (1979, p. 196).

In order to arrive at a better understanding of the tonal and temporal interplay it is necessary to isolate the critical points of an Fo-contour and to examine their variability in time and frequency. In spite of the relative inertia of the laryngeal system compared to the supra-laryngeal system Fo appears to be relatively flexible in the execution of a complex tonal pattern. Interruption, anticipation and delay of Fo-changes are possible ways of coping with this kind of situation (see Bruce 1977, chapter 5). For example the terminal juncture fall occurs with a delay in the accent 2-word `nunnor in focus (Figs. 2 and 3, upper part) relative to the accent 1-word ´nummer. This flexibility of Fo makes it difficult to see how tonal demands could affect the segment durations to any considerable degree.

The results from the present study of tonal and temporal interplay in Standard Swedish indicate that there is a basic, temporal patterning which is independent of the tonal patterning. It has not been possible to confirm that this basic, temporal pattern is modified by certain tonal demands.

REFERENCES

Bannert R. 1979. "The effect of sentence accent on quantity". Proceedings from the ninth international congress of phonetic sciences. Volume II: 253-259. Institute of Phonetics. University of Copenhagen

Bruce G. 1977. Swedish word accents in sentence perspective. Lund: Gleerup

Bruce G. and E. Gårding. 1978. "A prosodic typology for Swedish dialects". Nordic Prosody, E. Gårding et al. (eds): 219-228. Department of Linguistics. Lund University

Erikson Y. 1973. "Preliminary evidence of syllable locked temporal control of Fo". STL-QPSR 2/3: 23-30

Lindblom B., B. Lyberg and K. Holmgren. 1976. Durational patterns of Swedish phonology: Do they reflect short-term motor memory processes. Institute of Linguistics. University of Stockholm. (Mimeographed)

Lyberg B. 1979. "Final lengthening - partly a consequence of restrictions on the speed of fundamental frequency change?" Journal of Phonetics 7: 187-196

Öhman S., S. Zetterlund, L. Nordstrand and O. Engstrand. 1979. "Predicting segment durations in terms of a gesture theory of speech production". Proceedings from the ninth international congress of phonetic sciences. Volume II: 305-311. Institute of Phonetics. University of Copenhagen