

THE LINGUISTIC FUNCTIONS OF F0 PEAKS

KLAUS J. KOHLER

Institut für Phonetik und
digitale Sprachverarbeitung
Universität Kiel
2300 Kiel, FRG

ABSTRACT

F0 is an essential acoustic signalling property for both stress and intonation. Although sound duration and intensity are further characteristics of the former, a change in F0 may be sufficient to shift stress from one syllable to another. Taking a German minimal verb pair ("umlagern" with prefix or stem stress) as its point of departure, this paper presents experimental data showing (a) some of the conditions under which F0 is sufficient for a stress shift, and (b) the interaction of the stress and intonation functions of F0.

INTRODUCTION

It has been shown in /1,2/ that within the same word and sentence stress (e.g. the syllable "-lo-" in "Sie hat ja gelogen.") three types of intonation F0 peak positions are possible: early, medial, or late (in the syllable "ge-" or central or late in the syllable "-lo-" of the quoted example), with the corresponding changes of meaning from 'established' to 'new' to 'emphatic'. On the other hand, a shift of the F0 peak position from one syllable to another can change the stress position in a syllable chain. Thus two questions arise:

- (a) Under what conditions is an F0 peak shift (without concomitant changes in sound duration and intensity) sufficient to shift stress to a different syllable?
- (b) How can the stress and intonation functions of F0 peaks be differentiated, and in what ways do they interact?

To provide answers to these questions two experiments were carried out in German, which offers a good example for testing the issues because it has minimal verb pairs, with either prefix or stem stress, which can occur in the same natural sentence frame, e.g. "Er wird's wohl umlagern." (with stress either on "um-" /'um/ = "verlagern", "He is pre-

presumably going to shift it to another place."; or on "-la-" /'la:/ = "belagern", "He is presumably going to besiege it.").

PROCEDURE

Two utterances of this sentence, (a) with stress on "um-" and a medial intonation F0 peak on this syllable, and (b) with stress on "-la-" and an early intonation F0 peak, which actually falls on the syllable "um-", were selected for stimulus construction from a large corpus containing several repetitions of all the 6 combinations of 2 stress and 3 intonation positions, spoken by a trained phonetician (the author). The two tokens were analyzed using the same procedure as in /1/. Figures 1a,b present the waveforms together with their F0 displays. The F0 peak positions in the two utterances are practically identical in relation to the syllable structures of "umlagern": they occur at more or less the same time interval before the beginning of /l/. The differences between the two are in the shapes of the F0 peak contours and in the syllable durations:

- (a) in prefix stress, the F0 rise of the peak contour sets in at the beginning of "um-", in stem stress, however, as early as the beginning of /l/ in "wohl",
- (b) in prefix stress, "um-" is much longer than in stem stress (222 ms vs. 135 ms), but "-la-" has very similar durations in both cases (258 ms vs. 268 ms).

In a second step, the F0 peak contours of the two utterances were exchanged and adjusted to the comparable points in the segmental structures. Figures 2a,b show the waveforms of figures 1a,b with the new F0 contours. Finally, the following F0 parameter manipulations were performed:

- (1) In the stimulus of figure 1a (original prefix stress), the whole peak contour between the marks A and C was shifted to the right along the time axis in 6

equal steps of 30 ms; the tail of the F0 contour beyond mark C was then time-compressed between the new time position C' and the end of periodicity, and the F0 precursor in "wohl" was time-expanded from its beginning point to the new time position A'. The left branch of the peak contour (AB) was also shifted to the left in 5 equal steps of 30 ms; the right branch of the peak contour was then time-expanded between the new time position B' and the time mark C, and the precursor was time-compressed between its beginning and the new time position A'. When A' fell to the left of the beginning of "wohl" the section of the contour that thus entered the voiceless stretch was masked.

- (2) In the stimulus of figure 2b (original stem stress with transferred F0 peak shape), the whole peak contour between the marks A and C was shifted to the left in 8 equal steps of 30 ms; the tail of the F0 contour beyond mark C was then time-expanded between the new time position C' and the end of periodicity. As regards the left-branch adjustment the same procedure was followed as in the left shifts of (1).
- (3) In the stimulus of figure 2a (original prefix stress with transferred F0 peak shape), the same F0 peak shifts were carried out as in (1).
- (4) In the stimulus of figure 1b (original stem stress), the same procedure was followed as in (2).

From these parameter manipulations, there resulted 12 F0 contours, with peak positions from near the beginning of "um-" to the second half of "-la-", in (1) and (3), and 9 F0 contours, with peak positions from the beginning of "wohl" to near the end of "um-", in (2) and (4). These F0 contours entered into a stimulus synthesis with the LPC-derived formant and volume values of the original prefix-stress utterance in (1) and (3), and with the corresponding data of the original stem-stress utterance in (2) and (4). In each case, two test stimulus sets were thus generated, with a slowly and an abruptly falling F0 peak contour, respectively: (3), (4) vs. (1), (2). In (1) and (3), the F0 peak positions straddle the syllable structures where a change from prefix to stem stress is to be expected if F0 is a sufficient cue. The two sets differ in that the peak shape of (3), but not of (1), approximates the configuration found in the early peak of the original stem-stress utterance (see figure 1b). It is hypothesized, therefore, that if stress is perceptually shifted at all in (1) and (3), there will be a more clear-cut change in (1) because there is a higher probability in (3) that an F0 peak position on "um-" is not only perceived as a medial

peak with prefix stress but also as an early peak with stem stress. The same would apply to (4) as against (2).

To check these hypotheses two test tapes were compiled: (I) containing the 12 stimuli of (1) and the 9 of (2), (II) containing the 12 stimuli of (3) and the 9 of (4). (I) was produced in a short version with 5 repetitions of the 21 stimuli, and in a long version with 10 repetitions, with separate randomizations of the 105 and 210 test stimuli, respectively. (II) was only produced in a short version. Each stimulus sentence was preceded by a bleep and followed by a 4 s pause in which subjects were to answer, by ticking the appropriate boxes on prepared response sheets, whether the meaning of the perceived stimulus was "belagern" or "verlagern". 18 subjects did test (I) in its long version, 9 in its short one. 4 of the 18 deviated in their responses by judging the 9 stimuli of (2) exclusively as "verlagern". They were, therefore, dealt with separately and not included in figures 3 and 4. 16 subjects, some of whom had done test (I), took test (II) in later sessions. The subjects listened to the test tapes in several subgroups via a loudspeaker in a sound-treated room of the Kiel Phonetics Institute.

RESULTS AND DISCUSSION

Figures 3 and 4 present the results from these experiments for the 12-stimulus sets (1), (3) and for the 9-stimulus sets (2), (4), respectively. In the shift of the more sharply falling (original) F0 peak contour through the original prefix-stress utterance, there is a clear change from initial to stem stress, in spite of the duration of "um-" pointing to the former. F0 can thus override duration, particularly since the duration of the unstressed "-la-" syllable in the original utterance is very close to its duration under stress. In stimulus 10, which is the first in the ordering from 1 to 12 to yield an unequivocal stem-stress categorization with over 80% positive responses, the F0 peak position is 30 ms into the vowel of the syllable "-la-". This corresponds to the medial intonation peak on the stressed syllable found in /1,2/. The fact that the change from one stress category to the other is gradual rather than categorical can be related to some interaction of the stress and intonation functions of F0 because the more sharply falling F0 peak assumes positions before the beginning of the syllable nucleus /a:/ of "-la-" which can simultaneously function as the medial intonation peak in stressed "um-" and as the early intonation peak related to stressed "-la-". When the more slowly falling F0 peak is substituted the initial-stress category is not clearly

represented, the interference from the early intonation peak of the stem stress becomes too strong.

When an F0 peak contour is shifted through the original stem-stress utterance, there is no change between the stress categories: the answers are predominantly in favour of stem stress. In this case, F0 can thus not override the duration cue completely because "um-" is too short in relation to "-la-" to signal initial stress. But there is some effect of F0 when the more sharply falling F0 peak is moved into the syllable "um-": in stimulus 5 the peak is positioned at the end of /l/ in "wohl", with the F0 fall occurring in "um-", in all subsequent stimuli the peak is itself located in "um-". When the characteristic peak contour occurs in the relevant syllable the duration cue is checked by the F0 cue to a certain extent. The 4 subjects that behaved differently from the other 23 were guided by F0 altogether: they only perceived initial stress in all the 9 stimuli, where the F0 peak precedes "-la-". The substitution of the more slowly falling F0 peak reduces the prefix-stress judgements because of the interference of the intonation function of F0.

The hypotheses that led to the experiments discussed in this paper have thus been confirmed, and the questions asked initially can be answered as follows:

- (a) An F0 peak shift by itself is sufficient to bring about a clear change from one stress position to another, provided the duration of the stressed-syllable-to-be is not too short. But even when it is there is a residual F0 effect.
- (b) The intonation function of F0 interferes with its stress function if the latter is not supported by duration. This finds its expression in a gradual change from one stress category to another over a stretch of utterance where the positions of a medial intonation peak in one stressed syllable and an early intonation peak related to a stressed syllable following can coincide. This interaction is strengthened when the shape of the F0 peak contour approximates the more slowly falling one of the early intonation peak of a later stress.

REFERENCES

- /1/ K.J. Kohler, "Computer synthesis of intonation", Proc. 12th Intern. Congr. Acoustics, A6-6, Toronto, 1986.
- /2/ K.J. Kohler, "Categorical pitch perception", Proc. 11th Intern. Congr. Phon. Sc., Tallinn, 1987.

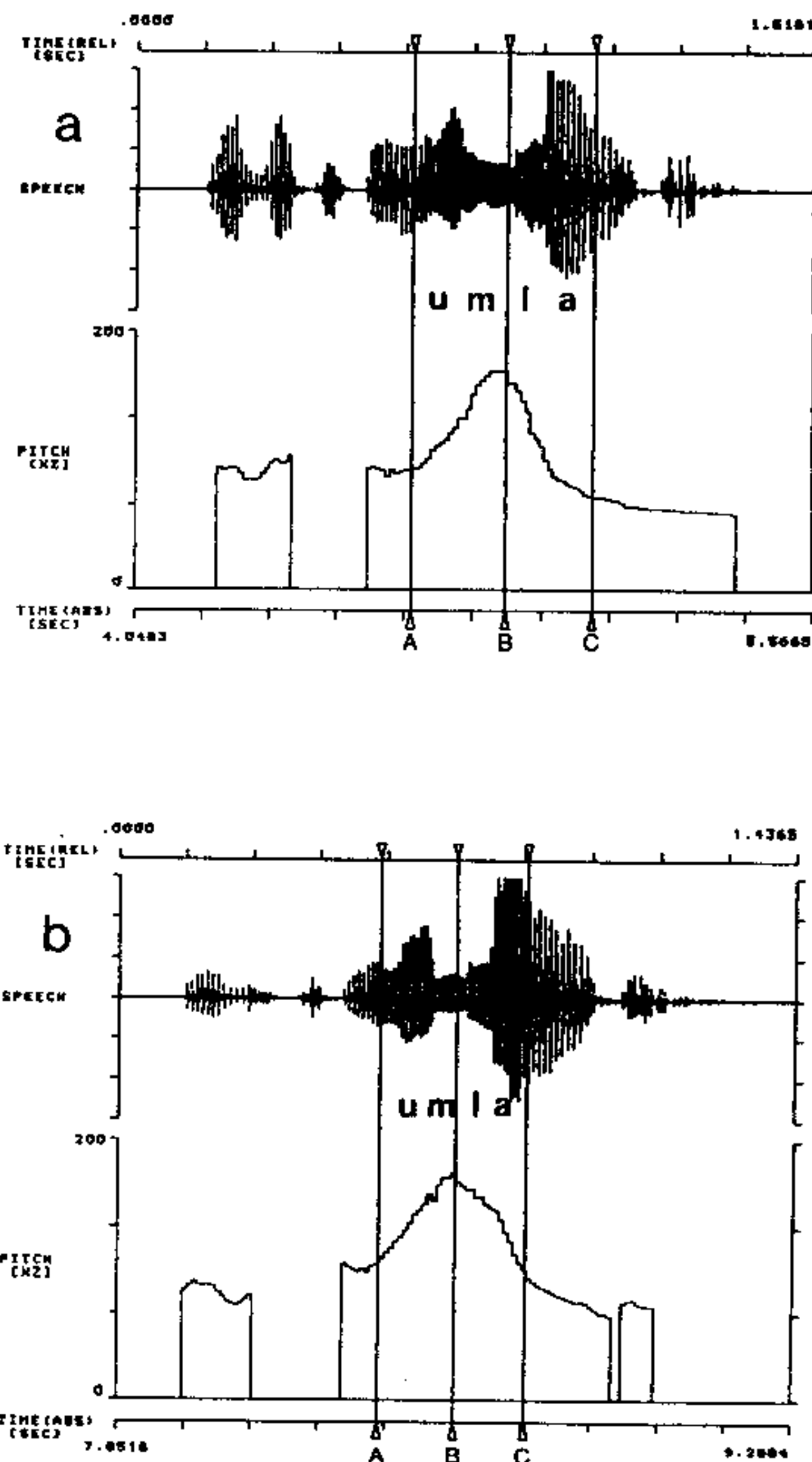


Fig. 1. Waveform and F0 of the original prefix-stress (a) and of the original stem-stress (b) utterance of "Er wird's wohl umlagern." A, B, C mark the base and peak points of the F0 peak contour for the F0 shifts.

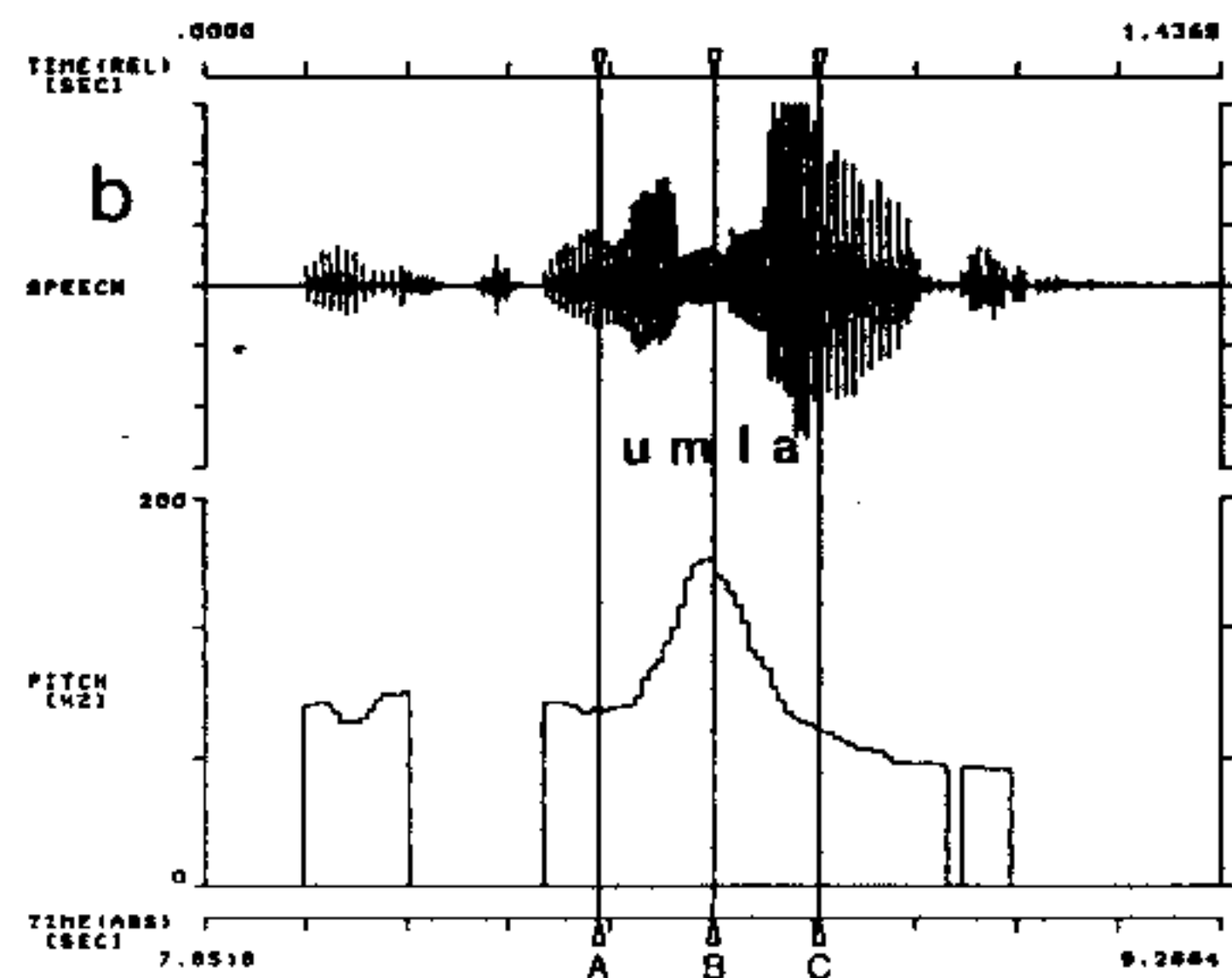
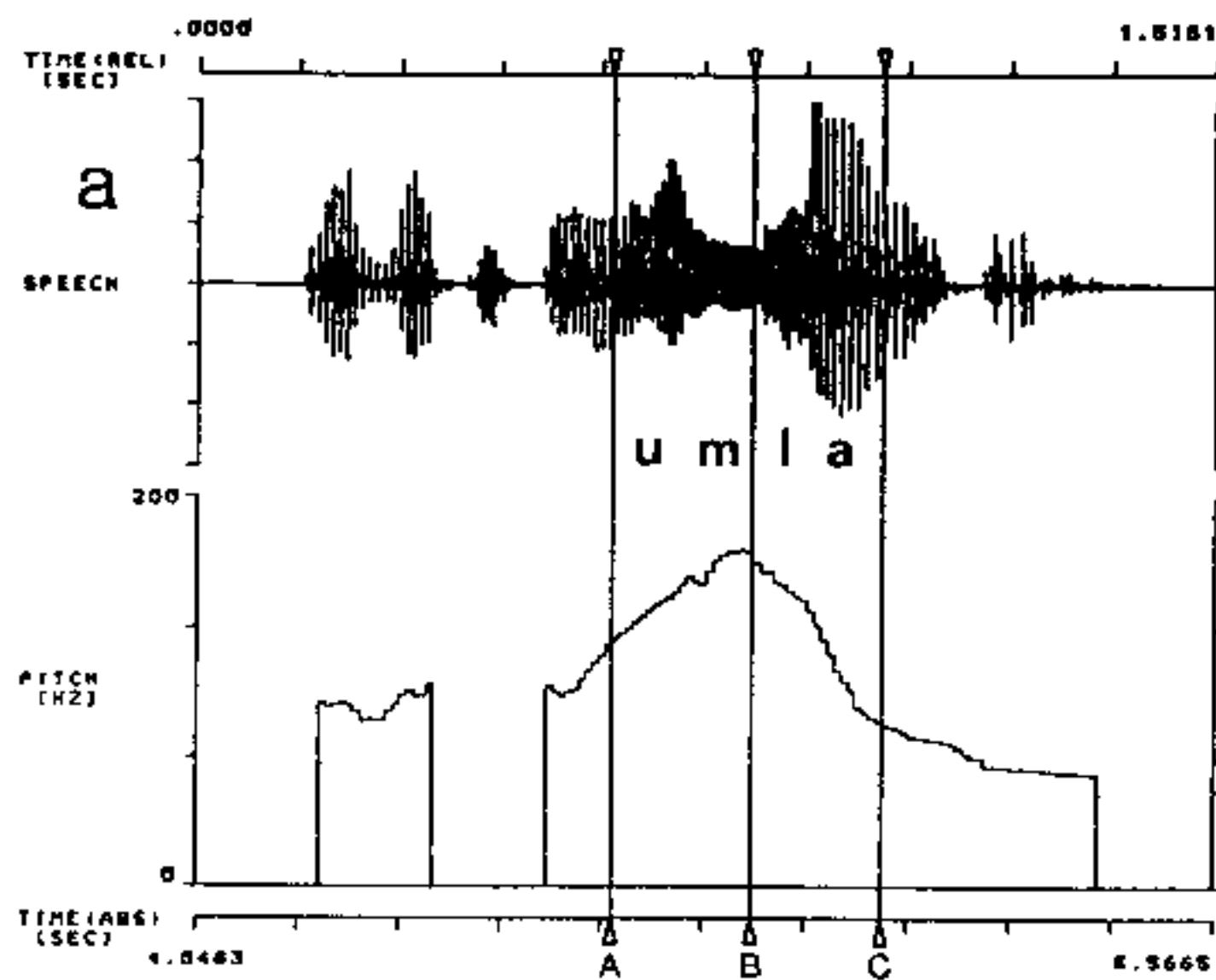


Fig. 2. Waveform of the original prefix-stress (a) and of the original stem-stress (b) utterance of "Er wird's wohl umlagern." with the F0 peak shape

transferred from the stem-stress (a) and from the prefix-stress (b) utterance and adjusted to the different timing of the new utterance. A, B, C as in figure 1.

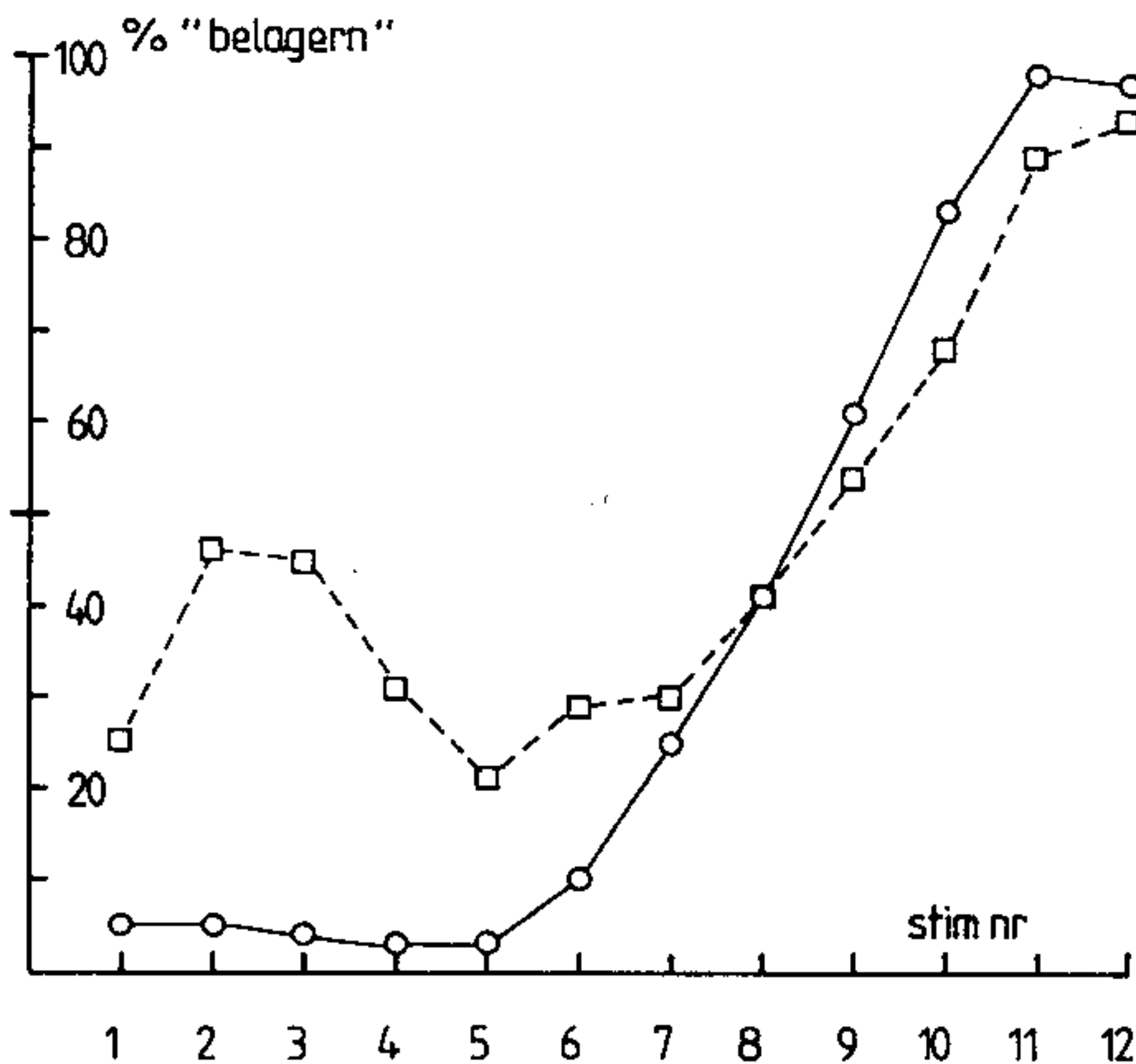


Fig. 3. Percentage stem-stress responses for "umlagern" (= "belagern") in the series of 12 F0 peak positions combined with the original prefix-stress utterance of "Er wird's wohl umlagern."; original, sharply falling peak contour (continuous line, at each data point $N=14 \times 10 + 9 \times 5 = 185$), and slowly falling peak contour, transferred from the original stem-stress utterance (broken line, at each data point $N=16 \times 5 = 80$).

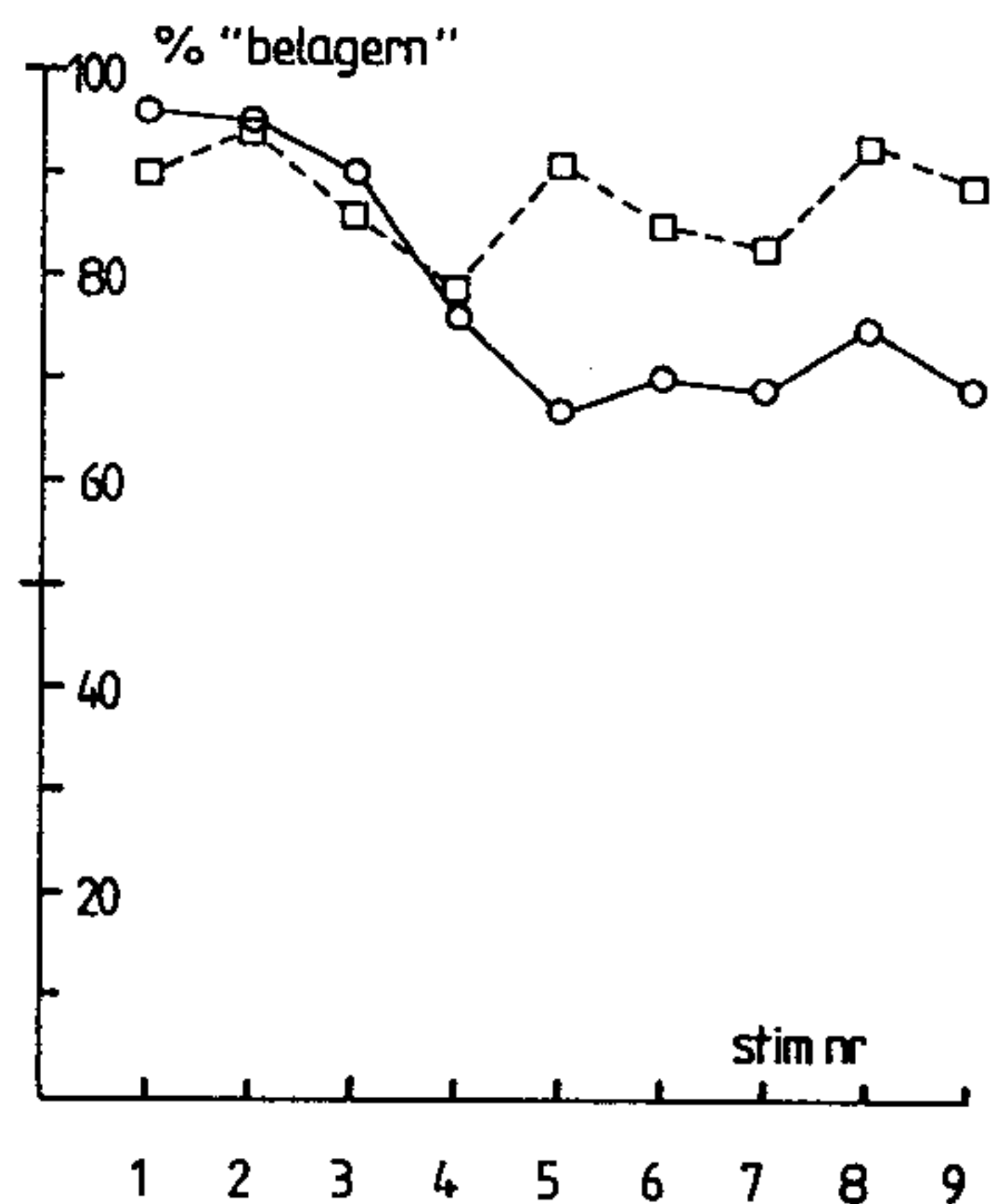


Fig. 4. Percentage stem-stress responses for "umlagern" (= "belagern") in the series of 9 F0 peak positions combined with the original stem-stress utterance of "Er wird's wohl umlagern."; original, slowly falling peak contour (broken line, at each data point $N=16 \times 5 = 80$), and sharply falling peak contour, transferred from the original prefix-stress utterance (continuous line, at each data point $N=14 \times 10 + 9 \times 5 = 185$).