ASSIMILATION OF Fo

Niels Reinholt Petersen
Institute of Phonetics, University of Copenhagen

1 Introduction

The observed fundamental frequency course of an utterance can be viewed as the result of contributions from several simultaneous components. In Advanced Standard Copenhagen (ASC) Danish four such components have to be considered (see Thorsen, 1979): (1) a sentence component, which gives the intonation contour of the sentence, (2) a stress group component, which supplies the Fo movements of the stress groups (a stress group in ASC Danish is constituted by a stressed syllable plus the following unstressed ones), (3) - in words with 'stød' - a stød component, and (4) a segmental (or microprosodic) component, which gives the Fo variation attributable to the segments constituting the utterance, such as inherent Fo differences and coarticulatory effects on Fo across segment boundaries.

Since it is a consequence of inherent properties of the speech production system and cannot be voluntarily controlled by the speaker, the segmentally determined Fo variation cannot - unlike the variation accounted for by components 1, 2, and 3 - carry linguistically relevant prosodic information. On the contrary, if sufficiently large the segmentally determined Fo variation could be expected to interfere with and possibly distort the contributions of the linguistically relevant components. The present paper is intended to examine the interaction between the segmental component (with the emphasis on the inherent Fo differences between vowels) and one of the linguistically relevant components, viz. the stress group component.

The Fo pattern of the ASC stress group can be described as a relatively low stressed syllable followed by a high falling tail of unstressed syllables (Thorsen, 1979). The Fo rise from the stressed to the first post-tonic syllable varies between 3 and 0.5 semitones. These values are very similar to the inherent Fo differences between high and low vowels, which have been found by Reinholt Petersen (1978, 1979) to vary between 1 and 3 semitones.

In theory this might imply that the Fo contour of an utterance could be distorted in such a manner that the perceived stress pattern would differ from that intended by the speaker as a function of the qualities of the vowels in the utterance. Now, such things do not happen; people normally hear stress patterns as intended by the speaker. One explanation for this could be that the listener perceptually compensates for inherent Fo differences, and thus reconstructs the intended Fo contour. However, the inherent Fo differences are considerably larger in stressed than in unstressed syllables (Reinholt Petersen, 1979), and this means that the perceptual system would have to know the stress pattern in order to be able to select the appropriate correction factor for the reconstruction of the intended stress group contour. Another possibility, which is the one under investigation here, could be hypothesized, namely that the compensation takes place in the speech production system as a coarticulatory assimilation of Fo between syllables which more or less smoothes out the distortion and thereby preserves the intended Fo contour.

2 Experiment I

In experiment I the following questions were considered: (1) Can Fo in one syllable be influenced by high vs. low vowel (i.e. a vowel with high vs. low inherent Fo) in adjacent syllables? (2) Is the effect, if any, directional (i.e. is it the preceding or the following vowel that has the stronger effect)? (3) Are there differences between syllables in different positions in the stress group as to how strongly Fo is influenced?

2.1 Method

The material consisted of mVmVmVmV nonsense words in which the vowels i and ɑ and u and ɑ alternated in such a manner that it was possible to see how Fo in these vowels was influenced by i vs. ɑ and u vs. ɑ in preceding and following syllables in the stress group positions 1st pretonic (or more correctly, in the last unstressed syllable in the preceding stress group), stressed, 1st post-tonic, and 2nd post-tonic. The words - embedded in frame sentences - were read by two female (KM, SI) and two male (PA, NR) ASC speakers. Six repetitions of each word were obtained for each speaker.

2.2 Results

Fig. 1 shows the average Fo in semitones over all subjects as a function of high vs. low vowel in the preceding and the following syllables in the four stress group positions under investigation. It is seen from fig. 1 that Fo is fairly consistently higher after i and u than after ɑ. The strongest effect is found in the first post-tonic syllable, where it approaches one semitone. The following vowel seems to have no consistent effect on Fo. Thus the present data may be taken as evidence for a progressive assimilation of Fo which tends to preserve the intended Fo contour, and particularly the Fo rise from the stressed to the first post-tonic syllable, although - it must be emphasized - the inherent difference of about two semitones in the

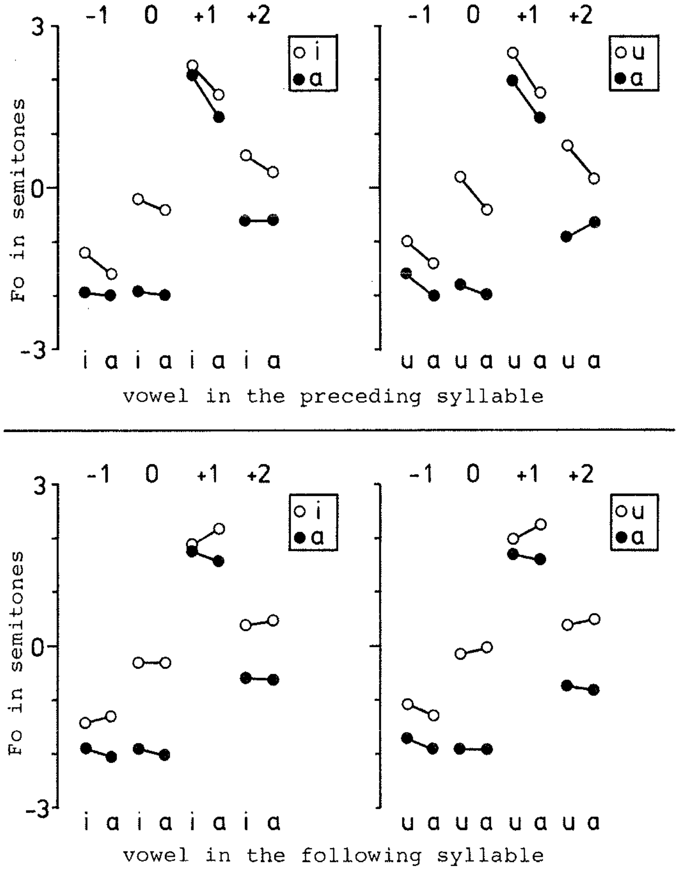

Figure 1. Effect of high vs. low vowel in preceding (upper graph) and following (lower graph) syllable on Fo in 1st pretonic (-1), stressed (0), 1st post-tonic (+1), and 2nd posttonic (+2) syllable.

stressed syllable is compensated by less than one semitone in the first post-tonic.

3 Experiment II

The aim of experiment II was to see (1) whether an intrinsic Fo difference in the stressed syllable ascribed not to high vs. low vowel but to the initial consonant would influence Fo in the first post-tonic, and (2) if the effect of high vs. low vowel in the stressed syllable on Fo in the first post-tonic would be the same whether the consonant intervening between the two syllables was voiced (as in experiment I) or voiceless.

3.1 Method

The material consisted of nonsense words representing all 'CVCV combinations of C = m and f and V = i and ɑ. The words were inserted in carrier phrases and read by the same subjects as in experiment I. Six repetitions of each word were obtained per subject.

3.2 Results

A prerequisite for considering at all question (1) above was that initial f vs. m had an effect on Fo in the same syllable. In the present material Fo was found to be 0.9 semitones higher after f than after m in stressed syllables. This figure is comparable in magnitude to the Fo difference between i and ɑ, which was found to be 1.3 semitones in this material. On this basis the effects of i vs. ɑ and of f vs. m in the stressed syllable on Fo in the first post-tonic might also be expected to be of comparable magnitudes. This is not the case, however. From fig. 2 it appears that whereas Fo in the first post-tonic is higher

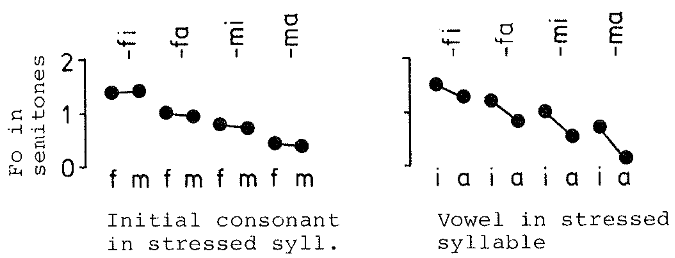

Figure 2. Effect of initial f vs. m (left) and i vs. ɑ (right) in the stressed syllable on Fo in first post-tonic -fi, -fɑ, -mi, and -mɑ syllables.

after i than after ɑ in the stressed syllable, the initial consonant in that syllable has only a very slight effect, if any at all, on Fo in the first post-tonic.

The effect of i vs. ɑ in the stressed syllable on Fo in the first post-tonic is present whether the consonant intervening between the two syllables is voiced or voiceless, although the effect seems to be slightly smaller across f than across m.

4 Discussion

As shown in the experiments reported above the effect of high vs. low vowel in the stressed syllable on Fo in the first post-tonic approaches one semitone. If an effect of that magnitude were to be accounted for by coarticulatory assimilation of tongue height in the first post-tonic vowel, a radical quality shift should be expected in that vowel as a function of the tongue height of the vowel in the stressed syllable. This applies whether the inherent Fo differences between vowels are to be explained by an acoustic source/tract coupling hypothesis or by a physiologically based tongue pull hypothesis (a review of these hypotheses is given eg. in Ohala, 1973). However, no such change is immediately audible, and measurements of F1 and F2 in ɑ in first post-tonic syllables of the material of experiment I show only a moderate effect on the formant frequencies from vowels in adjacent syllables (see fig. 3). Furthermore, the formant

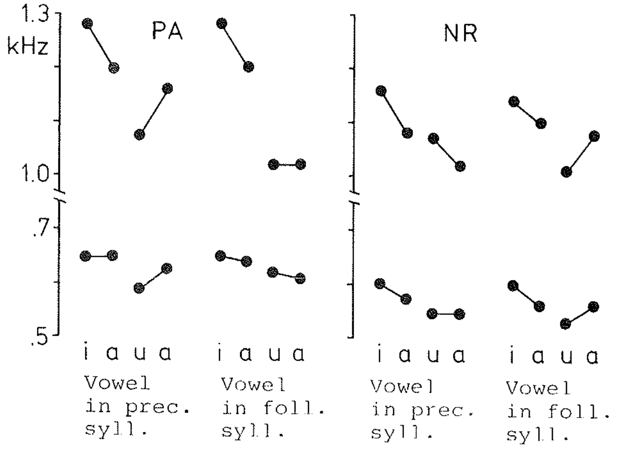

Figure 3. F1 and F2 in ɑ in first post-tonic syllables in the material of experiment I as a function of high vs. low vowel in the preceding and following syllables. Subjects PA and NR.

frequencies (particularly the second formant frequency) seem to be rather systematically influenced by the vowels both in the preceding and the following syllables.

Thus, there seems to be a discrepancy between the effect of neighbouring vowels on Fo and the effect of neighbouring vowels on the position of the tongue body, both as regards magnitude and direction. (This statement presupposes, of course, that formant frequency changes in the present material can be ascribed to tongue body position changes, which can be taken to be the case in the i/ɑ words, but not with safety in the u/ɑ words, where lip rounding is involved).

The discrepancy may be explained, however, if - under the tongue pull hypothesis - the relation between tongue height and vocal cord tension can be described in terms of a spring-mass system, where the driving force is the tongue body, the spring represents the connection between the tongue body and the laryngeal structures, and the mass represents vocal cord tension. If this description is tenable, changes in vocal cord tension, and hence in Fo, will occur with a delay relative to the changes in tongue height which produce them.

References

Ohala J.J. 1973. Explanations for the intrinsic pitch of vowels. Monthly Internal Memorandum, Phonology Laboratory University of California, Berkeley, 9-26

Reinholt Petersen N. 1978. Intrinsic fundamental frequency of Danish vowels. Journal of Phonetics 6, 177-189

Reinholt Petersen N. 1979. Variation in the inherent Fo level differences between vowels as a function of position in the utterance and in the stress group. Annual Report of the Institute of Phonetics University of Copenhagen 13, 27-57

Thorsen N. 1979. Lexical stress, emphasis for contrast, and sentence intonation in Advanced Standard Copenhagen Danish. Proceedings of the 9th International Congress of Phonetic Sciences. Vol 2, 417-423. Copenhagen