THE ROLE OF INTRINSIC FUNDAMENTAL FREQUENCY IN THE PERCEPTION OF SINGING

Niels Reinholt Petersen
Institute of Phonetics
University of Copenhagen

Introduction

It is a well established fact that in speech high vowels have higher fundamental frequencies (Fo) than low vowels, everything else being equal. Although the mechanisms underlying these intrinsic Fo differences among vowels are still subject to discussion, it is generally aggreed that, in contradistinction to the Fo variation associated with higher level prosodic categories (such as stress, tone, intonation), they are not under the voluntary control of the speaker.

In singing where a high degree of precision of Fo control is crucial for the realization of musical categories intrinsic Fo differences between high and low vowels are also observed, although they are smaller than those found in speech. In a study by Greiffenberg and Reinholt Petersen (1982) on unaccompanied solo-singing, two trained singers were instructed to sing an interval of an ascending minor third (embedded in a short carrier tune) on the vowel sequences *a-a, a-u, u-a,* and *u-u.* As it turned out Fo in the vowel *a* was lower than in *u* by 40 cents on the average (range 21-50 cents) at the same intended pitch, and - as a consequence - the minor third was produced almost one semitone wider in the *a-u* vowel sequence than in the *u-a* sequence (see figure 1). The two singers also read the text of the song, and here the intrinsic Fo differences were far larger than in singing, averaging 320 cents in stressed and 75 cents in unstressed syllables.

Intrinsic Fo differences in singing have also been also reported by Ternström, Sundberg, and Colldén (1983 and 1987) who had singers sustain a tone at a given pitch and make a change of vowel quality at mid-tone. A change between high and low vowels produced a fundamental frequency change approaching one semitone in a great majority of cases, when the singers' auditory feed back was masked by noise. With normal auditory feed back the Fo changes were smaller, but equally frequent.

From the results outlined above it is evident that the effect of vowel height on Fo is reduced in singing as compared to speech, but it is not eliminated and it is still well above the threshold of detection. On this basis the question arises why the singers listening to their own voices did not detect their deviations from the intended musical intervals or their fundamental frequency changes in the sustained notes.

The experiment reported below was intended to try out the possible explanation that in unaccompanied solo singing (choir-singing and accompanied singing may be different) listeners, including singers listening to their own voices during singing, perceptually compensate for the intrinsic Fo differences among vowels, i.e. the perceptual system expects, as it were, the same note to be realised at slightly different fundamental frequencies depending on the vowels sung.

Method

The test stimuli were synthetically produced tunes consisting of four *dV* -syllables. There were four different vowel sequence conditions, namely *dadadada, dadadadu, dudududa,* and *dudududu.* The first three syllables all had a fundamental frequency of 110 Hz (A2). In the fourth syllable Fo was varied in 11 steps from 124 to 139 Hz, i.e. from 207 to 405 cents above the first three syllables. The stimuli are schematically displayed in figure 2. The steps were

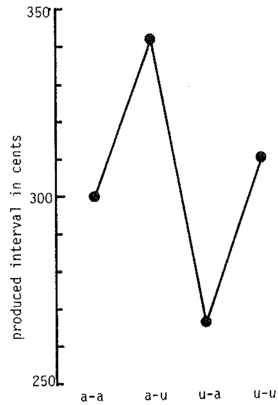

Figure 1: An intended minor third as produced under different vowel sequence conditions (adapted from Greiffenberg and Reinholt Petersen 1982).

intended to be 20 cents, but had to be rounded to the nearest integer number of Hertz, resulting in unequal step sizes. The eleven Fo steps were combined with each of the four vowel sequence conditions, making altogether 44 stimuli which were arranged in 10 different random orders in a test tape.

The test was taken individually by two musically trained female listeners, both of whom were university students of musicology and phonetics as well as experienced singers. The task of the listeners was to decide for each stimulus whether the musical interval between the third and the fourth syllable in the tune was a *major second,* a *minor third,* or a *major third.*

Results

The response patterns of the two listeners were very similar, and it was, therefore, thought jusitified to pool them. The identification functions for the three musical intervals are displayed in figure 2 (left graph), together with the 50% cross-over points (right graph).

It is clear that the boundaries are shifted systematically as a function of vowel height. The boundary between a major second and a minor third is about 20 cents higher in the *a-u* sequence than in the *u-a* sequence. The same pattern of variation is seen in the boundary between a minor and a major third, although the effect of vowel height is less pronounced. It is also seen from figure 2 that the median of the stimulus values having received minor-third-responses is influenced systematically by vowel height. An Extended Median Test showed the effect to be significant for the lower boundary ($p<0.001$) and for the median of the minor third responses ($p<0.05$), but not for the higher boundary ($p>0.05$).

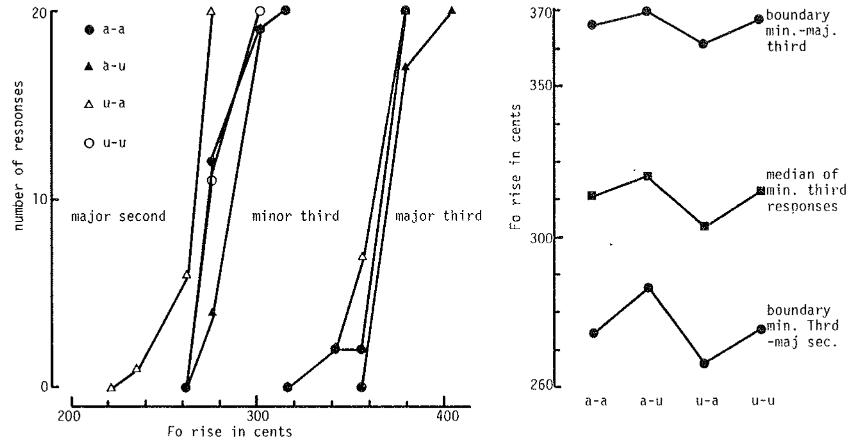

Figure 2: Identification functions (left), and 50% cross-over points and median values of minor-third-responses (right).

Discussion

The results of the experiment reported above has shown that the perception of musical intervals is systematically influenced by vowel height, and is thus giving support to the hypothesis that listeners to solo-singing perceptually compensate for vowel height. In addition to this main finding there are, however, two points which will also have to be considered in the discussion of the nature of the compensatory effect.

Firstly, the effect of vowel height revealed in the present experiment is far smaller than should be expected on the basis of the production data reported in Greiffenberg and Reinholt Petersen 1982 (cf. figure 1 above). The difference between extremes (i.e. a-u and u-a sequences) was about 80 cents in the production of an intended minor third, whereas the corresponding difference between medians of minor-third-responses in the listening test was only about 13 cents.

Secondly, as can be seen from figure 2 the effect is greater by a factor of more than two at the boundary between a major second and a minor third than at the boundary between a minor and a major third.

Research into the perception of pitch of vowels seems to give evidence for the existence of two types of vowel quality effect on perceived pitch. Although both are working in the direction to be expected under a hypothesis of perceptual compensation for intrinsic Fo, the two types of effect have - in my view - distinctly different origins (a more detailed treatment of this point may be found in Reinholt Petersen 1986).

One effect is a psychoacoustically conditioned *pitch bias effect* (as it has been aptly termed by Chuang and Wang 1978), and can be accounted for by reference to pitch perception processes common to the general class of complex periodic signals, of which vowels are, of course, a subset. The pitch bias effect is quite small, ranging from less than 1% to less than 3%.

The other type of vowel height effect on the perception pitch may be assumed to be speech specific, conditioned not by the spectral properties as such but by the phonetic characteristics of the vowels. It is demonstrated in experiments where listeners are to decide about prosodic categories of language (Rosenvold 1981, Silverman 1985), and - in contradistinction to the pitch bias effect - the magnitude of the phonetic compensation effect is very close to what should be expected from the intrinsic fundamental frequency differences found in speech production, i.e. about 10%.

In the perception of solo-singing it might very well be assumed that only the pitch bias effect and not the phonetic effect were involved. Actually, the relatively small boundary shifts observed in the present experiment point in that direction. This explanation does not, however, account for the fact that far larger deviations from an intended musical interval go unnoticed by singers listening to their own voices. Nor does it account for the shift being greater at the boundary between a major second and a minor third than at the boundary between a minor third and a major third.

On this basis it may be argued that the phonetic compensation effect is involved, after all, in the perception of solo-singing, but contributing to the perceived pitch to varying degrees depending on factors having to do with the individual listener, or with stimulus or listening conditions in general. Within this theoretical framework the discrepancy between the results of the Greiffenberg and Reinholt Petersen experiment and those of the present could be accounted for as follows: Listening to a human voice, and particularly, perhaps, to one's own voice during singing may be assumed to increase the degree of involvment of the phonetic effect in the perception of pitch, whereas listening to a synthetic voice will reduce it. Similarly it may be hypothesized that some musical intervals, and consequently their adjacent boundaries, are more resistant - so to speak - than others to the phonetic influence on perceived pitch, possibly depending (as was suggested to me by Robert McAllister, personal communication) on the functional importance they have in music. If this view is tenable, it could be the explanation for the the vowel height effect being less pronounced at the boundary between a minor and a major third that at the boundary between a major second and a minor third.

References

Chuang, C.-K. and Wang, W. S.-Y. 1978: Psychological pitch biases related to vowel quality, intensity difference, and sequential order, *J. acoust. Soc. Am. 64,* p. 1004-1014.

Greiffenberg, M. and Reinholt Petersen, N. 1982: The effect of high and low vowels on the fundamental frequency in singing: Some preliminary observations, *Ann. Rep. Inst. Phon., Univ. Copenhagen 16,* p. 101-111.

Reinholt Petersen, N. 1986: Perceptual compensation for for segmentally conditioned fundamental frequency perturbation, *Phonetica 43/1-3,* p. 31-42.

Rosenvold, E. 1981: The role of intrinsic Fo and duration in the perceptuion of stress, *Ann. Rep. Inst. Phon. Univ. Copenhagen 15,* p. 147-166.

Ternström, S., Sundberg, J., and Colldén, A. 1983: Articulatory perturbation of pitch in singers deprived of auditory feedback, *STL-QPSR 2-3,* p. 144-155.

Ternström, S., Sundberg, J., and Colldén, A. 1987: Articulatory perturbation of pitch in singers, *Rep. Uppsala Univ. Dept. of Linguistics 17,* p. 84-86.