Dept. for Speech, Music and Hearing
Quarterly Progress and Status Report

X-ray study of articulation and formant frequencies in two female singers

Johansson, C. and Sundberg, J. and Wilbrand, H.

KTH Computer Science and Communication

journal: STL-QPSR
volume: 23
number: 4
year: 1982
pages: 117-134

B. X-RAY STUDY OF ARTICULATION AND FORMANT FREQUENCIES IN TWO FEMALE SINGERS

C. Johansson,* J. Sundberg, and H. Wilbrand**

Abstract

Articulation and formant frequencies were studied from conventional radiographs of two professional female singers, an alto and a soprano. The subjects (1) pronounced the vowels [ɑ, i, u] in a speech mode, and (2) sang the same vowels at a low, a middle, and a high pitch. The jaw and lip openings, and, as regards the soprano, also the tongue shape and the larynx height changed systematically with fundamental frequency, in such a way that vowel differentiation mostly decreased with rising pitch. The frequency of the first formant was found to be close to that of the fundamental in cases where the fundamental would otherwise be higher than the first formant. As expected from the articulatory data, the formant frequencies showed a trend to increasing neutralization with rising pitch.

Introduction

Comparatively little is known about the characteristics of the female voice as compared with the male voice. The background is the high fundamental frequency range of the female voice which makes formant frequency estimates uncertain and, hence, information on the voice source unsafe. The situation is, of course, much worse in the case of singers than in the case of speakers, since singers typically use fundamental frequencies that are much higher than those used by speakers.

Some years ago one of the authors (JS) attempted to estimate the formant frequencies of a professional soprano singer for vowels sung at fundamental frequencies between 250 Hz and 700 Hz, approximately (Sundberg, 1975). These estimates were based on data obtained from external vibration of the vocal tract during "silent singing" as well as from matchings of the radiated spectra on a formant synthesizer. The results suggested that the soprano tuned her first formant frequency to the vicinity of the fundamental frequency, as soon as otherwise the first formant would be lower than the fundamental. Also, the second, third, and fourth formant frequencies were observed to be changed in a pitch dependent manner. Photos of the lip opening of several female singers singing various vowels at different pitches supported the hypothesis that the jaw opening is an important articulatory parameter behind this tuning of the first formant frequency. The generalizability of the results was supported by the naturalness obtained in a synthesis of female singing which replicated the observed pitch dependent tuning of the formant frequency (cf. the sound illustration in Sundberg, 1977).

* Dept. of Musicology, Uppsala University
** Dept. of Diagnostic Radiology, University Hospital, Uppsala

However, the questions of (1) the details of articulation and (2) individual differences among singers with respect to articulation were left open. The aim of the present investigation was to supplement some articulatory data from soprano and an alto singer, and also to attempt to make an acoustical evaluation of these data.

Experiment

The subjects were two well known professional Swedish singers, one alto who sings at the Stockholm Opera and one soprano who earns her life as a concert singer.

The entire vocal tract including the contours of the lips, the hard and soft palate, the back pharynx wall, the tongue, and the glottis were reproduced by conventional radiography. The technique and equipment thereby used were identical to that described in a parallel investigation of "kölning", (see Johnson & al, this issue of STL-QPSR).

Both singers sang the vowels [u , ɑ , i] at three fundamental frequencies, about 150, 300, and 600 Hz for the alto, and about 230, 470, and 950 Hz for the soprano. Also, radiographs were taken when the singers sustained the same three vowels in a normal speech mode. Finally, one picture was taken during rest with the subject holding a ruler in front of her lips, so as to obtain the scale factor.

From these radiographs area functions were derived, and the formant frequencies associated with these area functions were determined by means of an acoustical model of the vocal tract. The procedure thereby used will be accounted for later.

Results

A: Physiological data

Several physiological data are available in a lateral radiograph of the vocal tract. In the present study only such articulatory variables have been examined that are known to be acoustically relevant.

The vertical position of the larynx was measured by means of the template shown in Fig. 1, where 0 represents the midpoint of the upper glottis contour during rest. Adapting this template on the other radiographs caused no problems, as both singers kept their heads in the same way when they sang the various notes as during rest.

Fig. 2 shows the data obtained. There is a clear difference between the subjects. For the soprano's lowest pitches, the larynx position is lower than what is observed in the spoken vowels. However, she raises

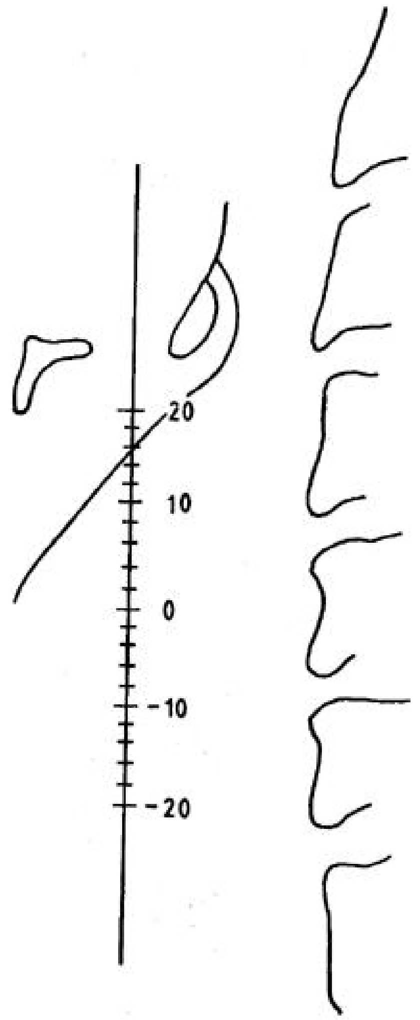

Fig. 1. Template used for measuring the vertical positioning in mm of the larynx. The value of 0 shows the rest position of the glottis contour.

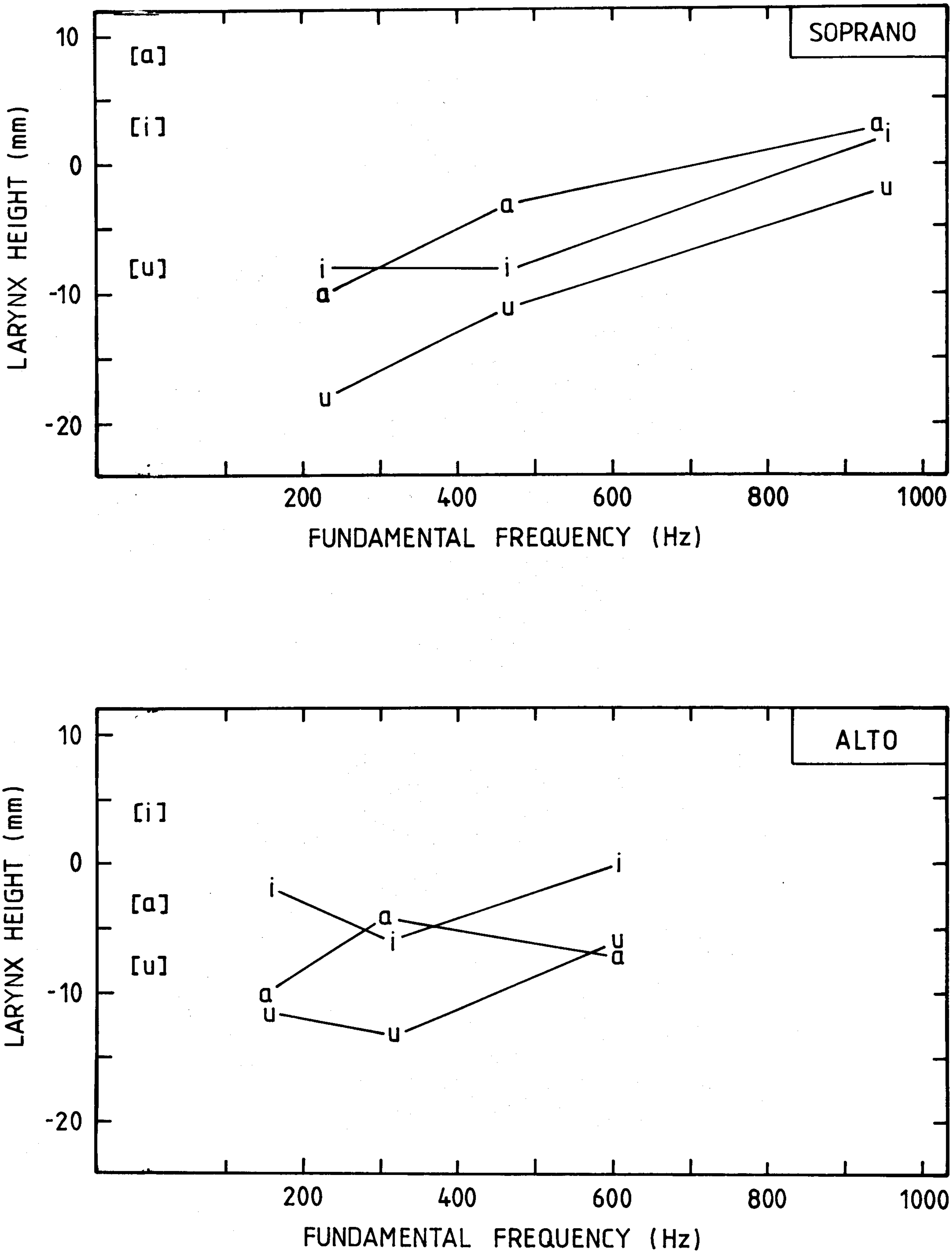

Fig. 2. The vertical position of the larynx as function of fundamental frequency in the two subjects. Vowel symbols within brackets pertain to spoken vowels, other symbols refer to sung vowels.

her larynx with fundamental frequency, so that she reaches a level similar to that of the spoken vowels at her top pitch. The alto uses a lower larynx in the sung vowels than in the spoken vowels in most cases but shows no clear relationship between larynx height and fundamental frequency.

In a previous investigation of larynx height in singing, Shipp and Izdebski (1975) found a lowering of the larynx that tended to increase with rising fundamental frequency in trained male singers. The data pertaining to our alto subject agree with this observation, while the opposite is true for the soprano. Therefore, the data were checked in the case of the soprano. Direct observation by means of fluoroscopy, when the soprano subject sang rising and falling triads, confirmed that she actually moved her larynx up and down in a harmoniously balanced synchrony with pitch.

The <u>jaw opening</u> was measured as the increase in distance between the upper and lower incisors beyond rest position with clinched teeth. The data are shown as function of fundamental frequency in Fig. 3. In this respect both subjects behave similarly in that they increase the jaw opening with rising fundamental frequency. Thus, their jaw opening values for the top pitches are far beyond what is observed in their corresponding spoken vowels. This is in agreement with previous observations (Ondráčková, 1969; Sundberg, 1975).

The lip opening is instrumental in tuning the formant frequencies. Two measures were taken from the lip contours. The vertical distance between the upper and lower lip (<u>h</u> in Fig. 4) is shown as function of the jaw opening in Fig. 5. It can be seen that this distance is highly dependent on the jaw opening, the correlation coefficient (linear regression) being 0.94 and 0.79 for the alto and the soprano, respectively. The weaker correlation in the case of the soprano reflects the fact that she reduced the vertical distance between her lips for all[u]-vowels and increased it for all [i]-vowels except in the case of the top pitch. It seems that in this respect the alto adopts a more relaxed articulation than the soprano.

The distance between a line joining the frontmost contours of the lips and the contour of the lip corners was also determined (<u>l</u> in Fig. 4). When the radiograph showed two mouth corner contours because of parallax error, the average of the resulting two distances was used. The resulting measure, which we will call <u>mouth corner retraction</u>, is shown Fig. 6. Here again, the soprano seems to exert more articulatory

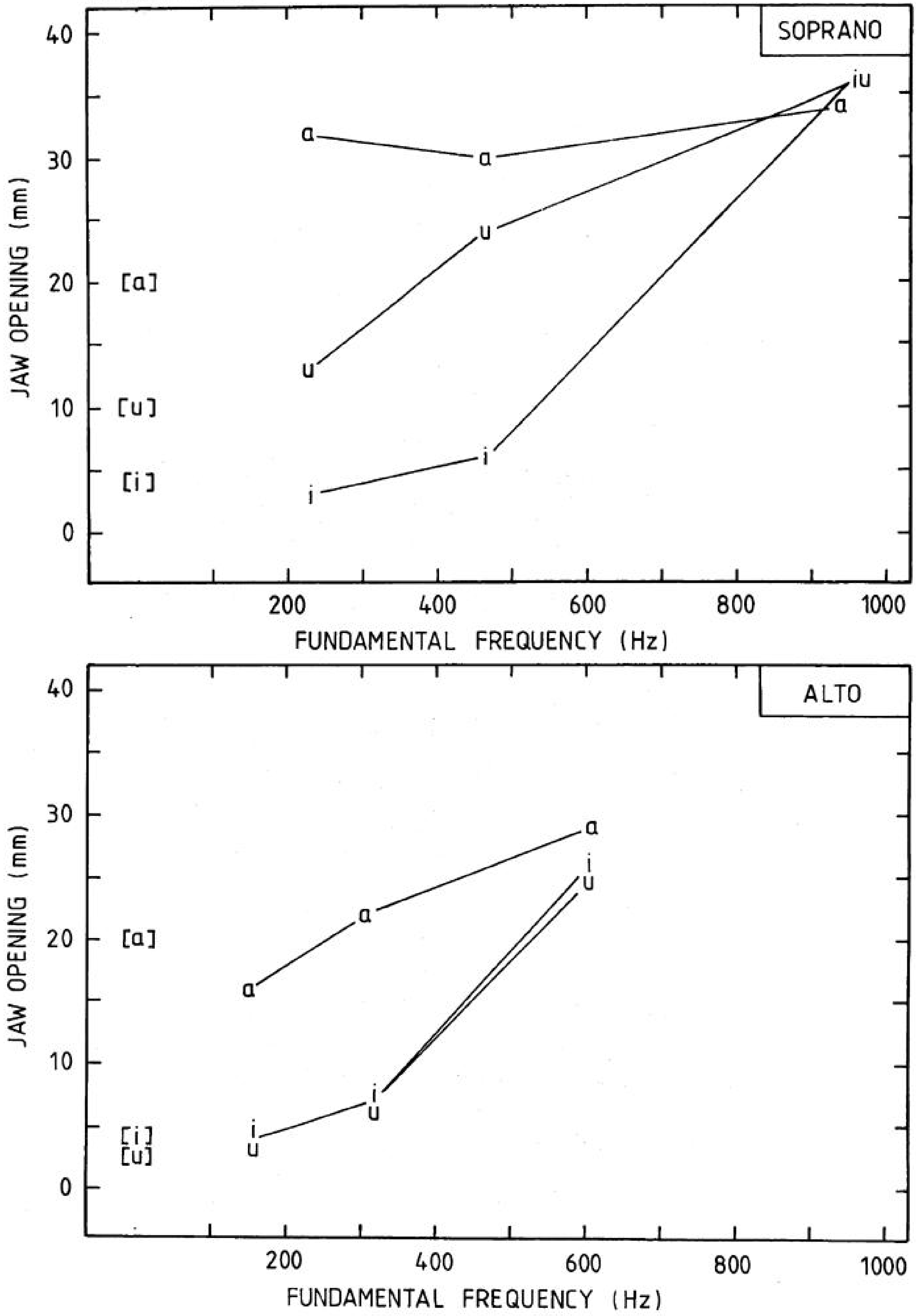

Fig. 3. The jaw opening as function of fundamental frequency in the two subjects. Vowel symbols within brackets denote spoken vowels, other symbols refer to sung vowels.

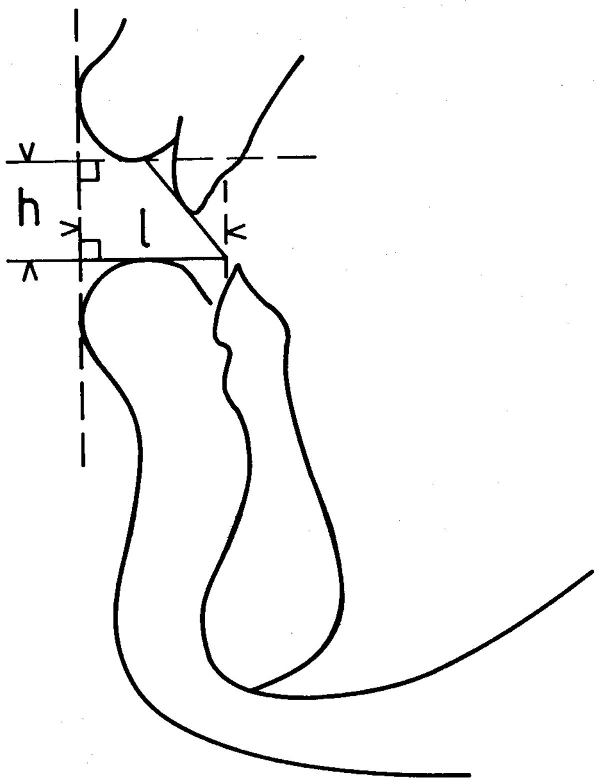

Fig. 4. Schematical illustration of the measurement of the vertical distance between the upper and lower lip, l, and of the retraction of the mouth corners, l.

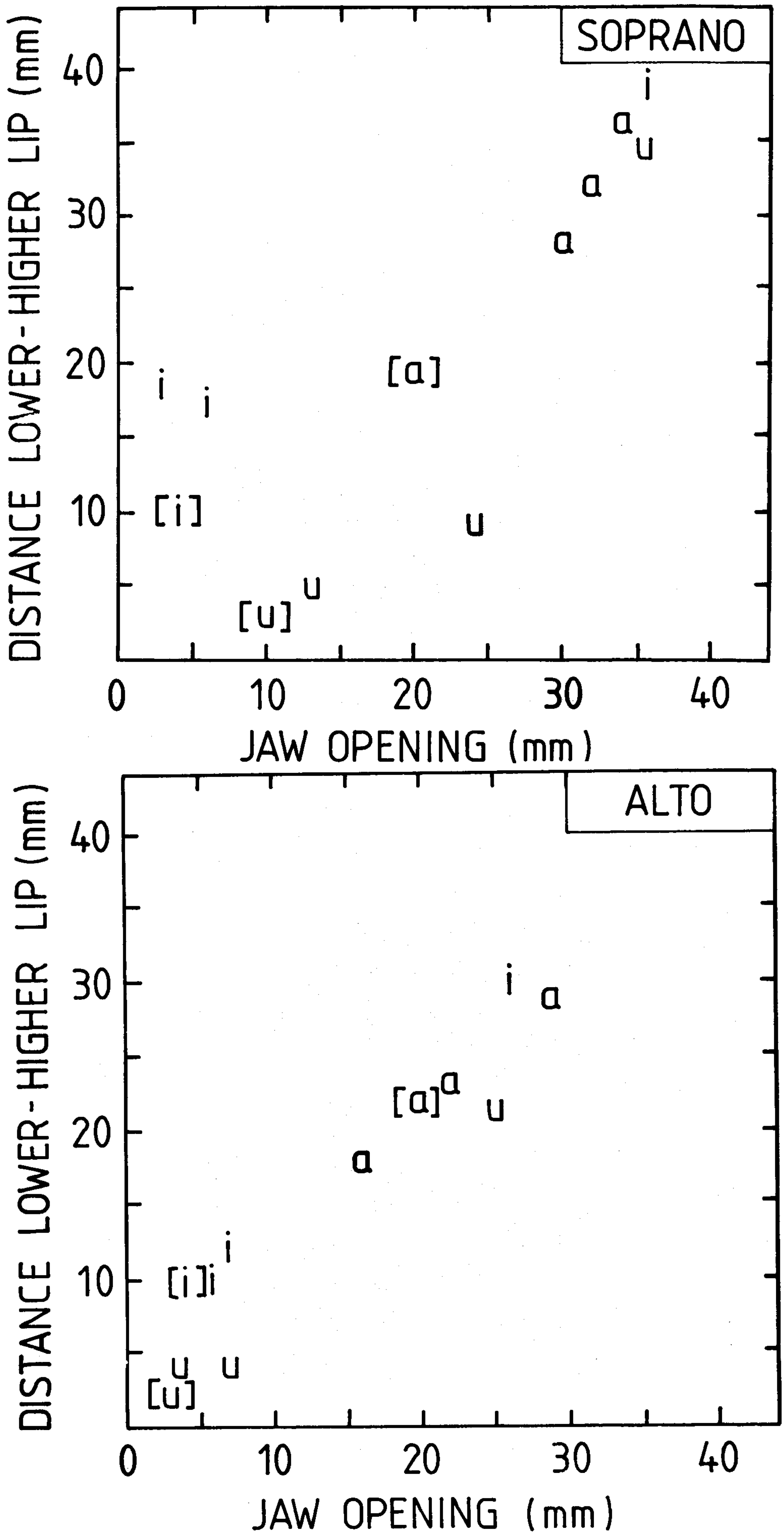

Fig. 5. The distance between the contours of the upper and lower lip as function of the jaw opening. Vowel symbols within brackers denote spoken vowels, other symbols refer to sung vowels.

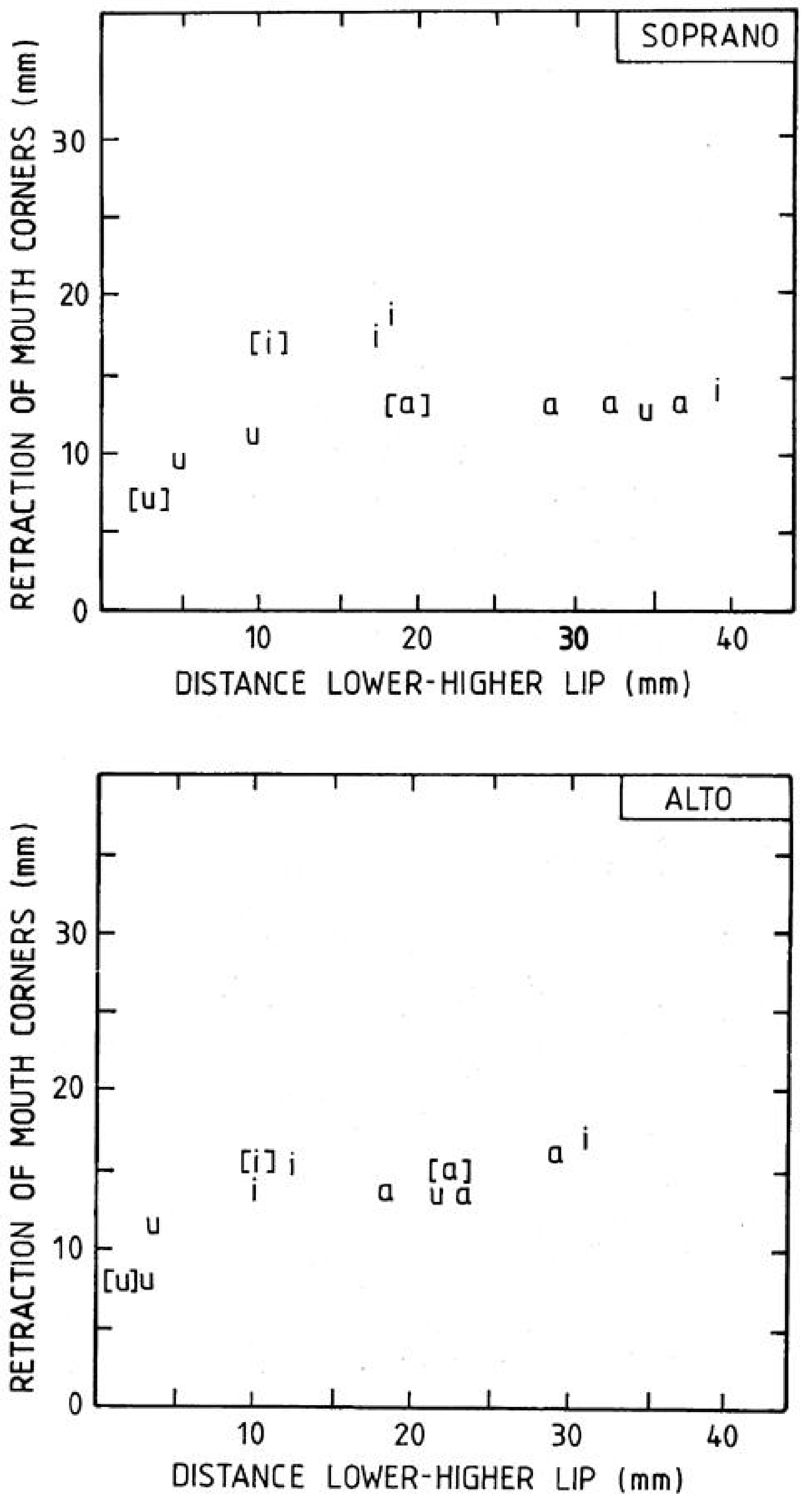

Fig. 6. Retraction of the mouth corners as function of the distance between the contours of the upper and lower lips. Vowel symbols within brackets denote spoken vowels, other symbols refer to sung vowels.

activity than the alto; the alto's data points are all close to one single curve, while the mouth corners seem to be actively retracted in the soprano's [i], again except for the highest pitch. In the case of the alto and in the case of the soprano's lower fundamental frequencies the three vowels show different values of mouth corner retraction, while these differences almost disappear in the soprano's highest fundamental frequency.

The tongue contours related to the lower jaw are shown in Fig. 7. The tongue contours are similar to those of the spoken vowels in the vowels sung at the lower fundamental frequencies. Also, both in the alto and in the soprano, the trend illustrated in Fig. 7a can be observed that the tongue shapes are slightly neutralized, as fundamental frequency is increased. At the soprano's highest pitch the tongue contours for all three vowels are similar to the tongue contour of the spoken [ɑ], as can be seen in Fig. 7b. Thus, while the alto uses at least two distinct tongue shapes for the three vowels at her highest fundamental frequency, the soprano uses practically identical tongue shapes for all vowels when she sings the highest note.

B: Acoustical data

Area functions were derived from the radiographs using the procedure described previously (Lindblom & Sundberg, 1971) and schematically illustrated in Fig. 8. First, points along the mid-line in the mid-sagittal plane of the vocal tract were determined. These points were derived as mid-points between the two intersections of the wall contours with the lines of grid of a semi-polar coordinate system. This system was anchored on contours of the hard palate and the cervical vertebrates, and was the same for all pictures. Then, these mid-points were smoothly joined and the resulting curve was regarded as a vocal tract mid-line. The distance between the vocal tract wall contours normal to this mid-line was determined at .5 cm intervals along the midline. The resulting mid-sagittal distances were then converted to cross-sectional area by means of two distance-to-area plots, one for the pharynx, and one for the mouth cavity. Lacking better data, the first-mentioned curve was the same as that used in a previous investigation (Sundberg, 1969; the possibilites of improving this procedure by means of recent development in radiology will be explored in a future investigation). The curve pertaining to the mouth cavity was derived from direct measurements on a plaster cast of the subjects' hard palates. Thus, for each radiograph an

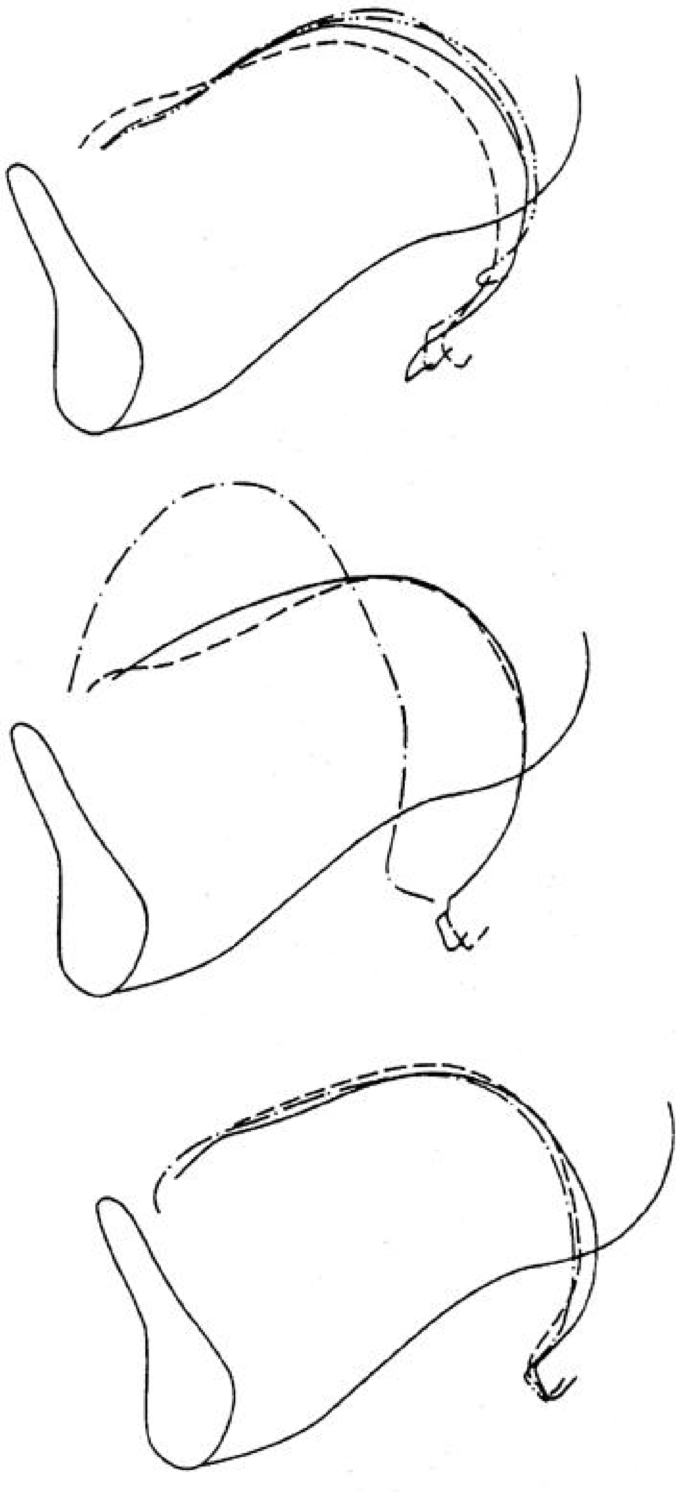

Fig. 7. Tracing of the midsagittal contours of the tongue body with the contour of the lower mandible as the reference. The upper group of tongue shapes are from the soprano's [ɑ] sung at the fundamental frequencies of 230, 465, and 940 Hz (solid, dotted-dashed, and dashed contours). The double-dotted-dashed curve is from the spoken version of the same vowel. The middle group of tongue contours are from the alto's [ɑ, i, u] vowels (solid, dashed, and chain-dashed curves) sung at 610 Hz fundamental frequency. The lowest group shows the corresponding tongue shapes for the soprano's [ɑ, u, i] (solid, dashed, and chain-dashed curves).

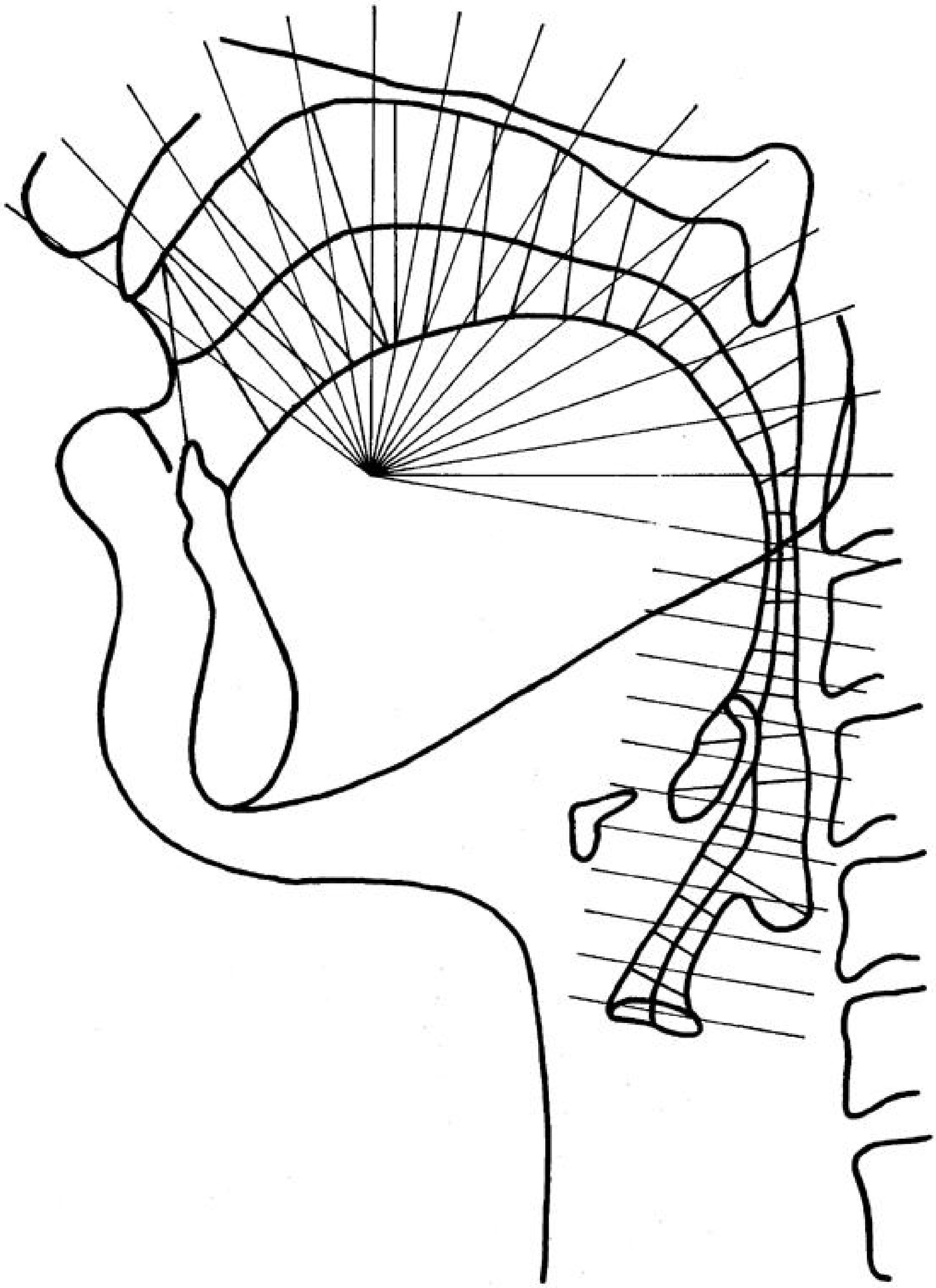

Fig. 8. Schematical illustration of the procedure used in order to estimate the dimensions of the vocal tract, see text.

area function was obtained, i.e., a table containing estimates of the vocal tract cross-sectional area at each .5 cm along the mid-sagittal vocal tract mid-line.

The next step was to determine the formant frequencies associated with each of these area functions. This was realized by piling plexiglass discs, .5 cm thick, each with a center hole of accurately determined size. By chosing discs with center holes in accordance with a specific area function, the resulting pile of discs constituted an acoustic model of this area function. These models were then excited by a sinewave of variable frequency by means of the STL Ionophone (Fransson & Jansson, 1973), so that the formant frequencies could be determined.

As is well known from acoustic theory of voice production, an increase of the jaw opening will raise the frequency of the first formant in particular, because it narrows the open-closed vocal tract tube in the closed end and expands it in the open end. Fig. 9 shows that this is the case with the formant frequencies obtained from the procedures just described, even though the soprano's spoken [ɑ] constitutes an exception. The reason for this is that the tongue shape of the spoken [ɑ] in itself constricts the pharynx and expands the mouth cavity, regardless of the jaw opening. Thus, the data given in Fig. 9 support the assumption that the formant frequency data represent reasonably reliable information.

The formant frequencies derived from the radiographs are shown in Fig. 10. It is evident that the data points pertaining to the three lowest formants form an organized pattern. This again supports the assumption that the formant frequency data are reasonably reliable.

For low fundamental frequencies the first formant is kept similar to what it is in the spoken version of the vowel. However, it tends to rise as soon as the fundamental frequency is raised above the frequency value of the spoken vowel. In other words, the situation seems to be avoided that the fundamental is higher than the first formant. At the top pitch, the frequency of the first formant is slightly lower than that of the fundamental, particularly in the case of the soprano. However, it has been shown that in reality the first formant frequency is raised by yielding vocal tract walls, and this effect increases with frequency. Thus, our first formant frequency estimates must be too low at high frequencies such as the soprano's top pitch. We conclude that, probably, the soprano and the alto both tune their first formant frequencies to the vicinity of the fundamental, as soon as the normal value

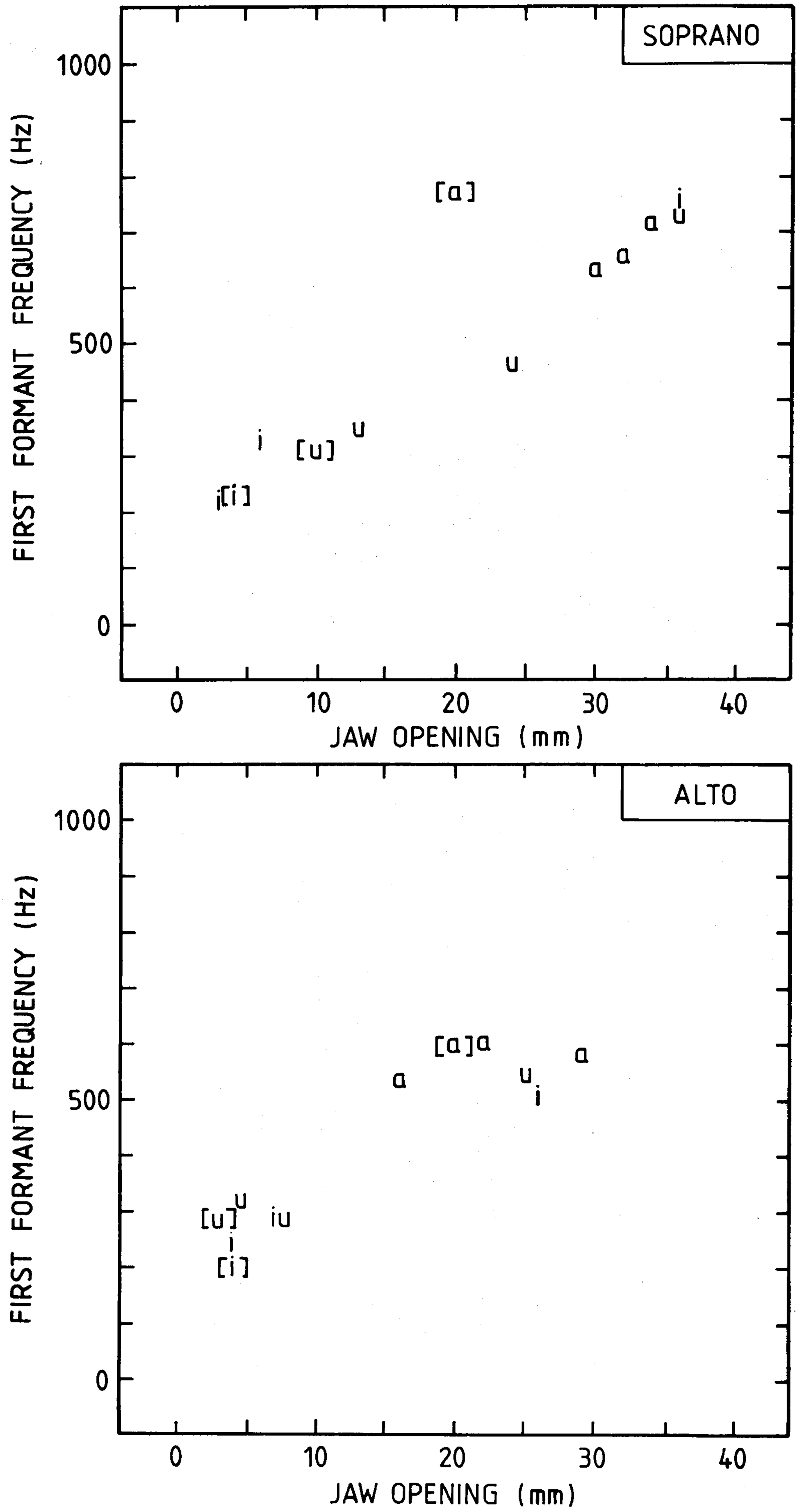

Fig. 9. Correlation between the frequency of the first formant and the jaw opening in the two singer subjects. Vowel symbols within brackets denote spoken vowels, other symbols refer to sung vowels.

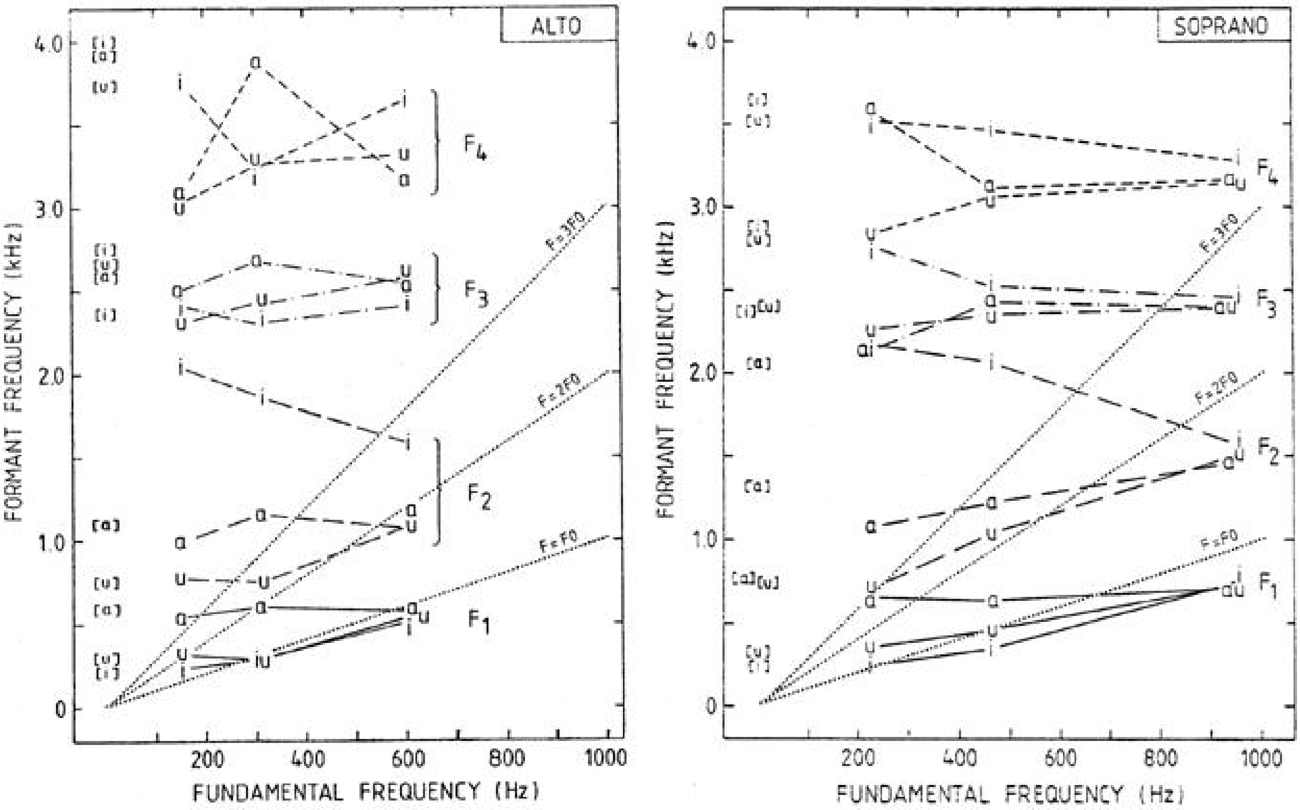

Fig. 10. Estimates of the frequencies of the first four formants as function of the fundamental frequency in the two singer subjects. Vowel symbols within brackets denote spoken vowels, other symbols refer to sung vowels. Symbols falling close to the dotted lines suggest that the formant frequency was tuned to the vicinity of the frequency of a harmonic partial.

of the first formant would imply that the fundamental was higher than the first formant.

With rising fundamental frequency, the second formant frequency rises in the back vowels [ɑ] and [u] and falls in the front vowel [i] in both singers. In the case of the soprano, this leads to the situation that the second formant frequency is practically identical for all three vowels at the top pitch.

The frequency of the third formant seems rather unaffected by pitch. The fourth formant frequency data seem unsystematic, even though they look strikingly similar to the larynx height data in the case of the alto.

Discussion and conslusions

With respect to the first and second formant frequencies of the various vowels sung at the different fundamental frequencies, our data are in good agreement with the results from the previous investigation of a soprano singer (Sundberg, 1975). This is interesting, given both the uncertainty of and the substantial method differences between the two investigations. Moreover, the subject used in the previous investigation was not identical with any of the subjects in the present study. Regarding the third and fourth formants, the results do not agree with those of the previous study. However, the data concerning these formants cannot be very accurate in the present investigation. The reason for this is that with rising frequency the formants become increasingly sensitive to details of the area function. For this reason these data do not merit any interpretation. We conclude that our data on the articulation and on the two lowest formant frequencies are reliable and that they possess a certain degree of generality.

The articulatory maneuvres underlying the tuning of formant frequencies involve jaw opening, larynx height, and tongue shape. The most important one would be the jaw opening, as both subjects were found to change this articulatory parameter with pitch; also, similar findings have been reported in previous studies of female singers (Ondráčková, 1969; Sundberg, 1975). The point of tuning the first formant to the vicinity of the fundamental in high pitched singing is a gain in SPL which may be quite considerable (cf., Sundberg, 1982). In high pitched singing it implies the need of raising the first formant to very high values.

The finding that the soprano raises her larynx with increasing fundamental frequency is interesting, since it is in clear opposition to the general opinion among many singing teachers, who claim that a rise of the larynx is harmful to the voice. Still, our soprano subject has been a successful professional solo singer for a number of years. We must conclude that an elevation of the larynx is not necessarily harmful to the voice (cf., Askenfelt & Sundberg, 1981).

The pitch dependent larynx height is also in clear opposition to previously published data regarding trained male singers (Shipp & Izdebski, 1975). However, from an acoustical point of view a soprano has good reasons for raising her larynx with fundamental frequency. Since she has to sing at very high fundamental frequencies, thereby preferably tuning her first formant to the vicinity of the fundamental, she has an extreme need of raising her first formant frequency. An elevation of the larynx shortens the vocal tract and thus adds to the possibilities of raising the frequency of the first formant.

As a consequence of the formant strategy observed, the acoustical vowel differentiation is reduced with rising pitch. Thus, particularly in the soprano's highest note, which, incidentally, is by no means the highest pitch that a soprano may have to sing, the formant frequencies of the three vowels become very similar. This is also in agreement with the articulatory data; with increasing pitch the articulatory similarity between the vowels increase, so that, in fact, the soprano uses almost the same articulation for all vowels in the case of the highest fundamental frequency (see Figs. 1, 2, and 3). It may be noted that, from the point of view of vowel quality perception, this strategy is harmless (Sundberg & Gauffin, 1982).

Even though there are reasons to assume that the data presented here represent reliable and representative information, it is evident that they do not allow any inferences regarding typical differences between alto and soprano voices. However, we would like to speculate that, apart from the obvious factor of vocal fold morphology, the pharynx length may be relevant. One support for this speculation is that the pharynx length is varied systematically in the soprano, viz., by means of larynx height. The typically "darker" vowel quality of an alto singer would arise from certain formant frequency combinations; the lower second formant in [i] and the lower first formant in [ɑ] would lead to a timbral difference of this type. These two formants in these specific vowels are both particularly sensitive to the dimensions of the pharynx. The difference in the larynx height data between the subjects may be of relevance in this connection. Also, typical differences in the higher formants are likely, which have not been revealed in the present investigation (see Dmitriev & Kiselev, 1979).

References

Askenfelt, A. & Sundberg, J. (1981): "Larynx height and voice source. A relationship?", STL-QPSR 2-3/1981, 23-36.

Dmitriev, L. & Kiselev, A. (1979): "Relationship between the formant structure of different types of singing voices and the dimensions of the supraglottal cavities", Folia Phoniatrica 31, 238-241.

Fant, G. (1960:) Acoustic Theory of Speech Production, Mouton, s'Gravenhage.

Fransson, F. & Jansson, E. (1973): "The STL-Ionophone: Transducer properties and construction", J. Acoust. Soc. Am. 58, 910-915.

Lindblom, B. & Sundberg, J.(1971): "Acoustical consequences of lip, tongue, jaw, and larynx movement, J. Acoust. Soc. Am. 50, 1166-1179.

Ondráčková, J. (1969): "Some remarks on the analysis of sung vowels, pp. 407-418 in Study of Sounds XIV, Phon. Soc. of Japan.

Shipp T. & Izdebski, K. (1975): "Vocal frequency and vertical larynx positioning by nonsingers and singers", J. Acoust. Soc. Am. 58, 1104-1106.

Sundberg, J. (1975): "Formant technique in a professional female singer", Acustica 32, 89-96.

Sundberg, J. (1982): "Perception of singing", pp. 59-98 in Psychology of Music, (D. Deutsch, ed.), Academic Press, New York.

Sundberg, J. & Gauffin, J. (1982): "Amplitude of the voice source fundamental and the intelligibility of super pitch vowels", pp. 223-228 in The Representation of Speech in the Peripheral Auditory System, (R. Carlson & B. Granström, eds.), Elsevier Biomedical Press, Amsterdam.