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

**Formant frequencies in choir
singers**

Ternström, S. and Sundberg, J.

journal: STL-QPSR
volume: 28
number: 4
year: 1987
pages: 043-055

**KTH Computer Science
and Communication**

<http://www.speech.kth.se/qpsr>

B. FORMANT FREQUENCIES IN CHOIR SINGERS

S. Ternström & J. Sundberg

Abstract

The four lowest formant frequencies are measured in eight members of the bass section of a good amateur choir under two conditions: (1) when reading the text of a poem aloud; and (2) when performing the same text as a song. Certain formant frequency differences are observed which are similar to those previously found between professional singers' spoken and sung vowels. The intersubject scatter of the three lowest formant frequencies was smaller, and the fourth formant was lower in singing.

Introduction

A primary esthetic concern of choir singing is that individual voices should not be discernible as such. Choir singers are required to achieve a blend of their own voice with the sound of the rest of the choir. In a general sense, this means that each singer, in addition to performing the music, should strive to make the sound of his or her own voice similar in character to that which is prevalent in the group.

Acoustically, this might imply a matching of loudness (Goodwin, 1980) as well as of formant structure and voice source properties (Rossing, Sundberg, & Ternström, 1986). In order to explain why a choir sounds the way it does, we clearly require knowledge of the vocal behavior of the singers. The present investigation compares the formant frequencies in speech to those in singing, as observed in eight bass choir singers. Comparisons are also made with solo singing, as measured in earlier experiments.

Previous work

A fair amount of work on the vowel articulation of singers has been reported (e.g., Sundberg, 1974; Rossing & al., 1986; 1987), especially concerning the "singer's formant" of professional opera singers. Sundberg (1974) found that professional singers used different formants in speech and in singing. In particular, the first formant frequency was found to be higher in most sung vowels, as compared to speech. Similarly, for front vowels, the second and third formants were considerably lower in singing. Furthermore, professional singers were found to cluster their higher formants, thereby generating the so-called

singer's formant. The unusual spectral energy distribution thus achieved makes the voice more penetrating and makes it easier for the solo singer to be heard, e.g., over a loud orchestra.

Cleveland (1976) studied the formant frequencies of male soloists of different vocal classifications, i.e., bass, baritone, and tenor. He found a small but systematic increase in all formants when going from basses to tenors.

Rossing & al. (1986; 1987) compared solistic and choral modes of singing. They found that singers, who were experienced both as soloists and choir singers, modified their voices somewhat when changing from a solo to a choir mode of singing. In the choir mode, the singer's formant was less prominent, and the amplitude of the fundamental was higher; hence, the articulation was closer to that of speech.

Goodwin (1980) studied how singers behaved when asked to blend with a choir they heard over earphones. He found that the subjects tended to use a softer voice level when blending and that their higher spectrum partials were relatively weaker, as might be expected, given the lower voice level. These investigations were mainly concerned with the personal voice timbre of individuals, being mainly dependent on the higher formants. For choirs, we also need to address the issue of vowel quality, being mainly determined by the two lowest formants. Thus, we might ask how the vowels are pronounced in choir singing as compared to speech or solo singing, and whether the singers in a choir do try to converge, e.g., for reasons of choral blend, on a particular set of formant frequencies for a given vowel. If so, do choir singers adjust to the vowel articulation of the ensemble momentarily, vowel for vowel, while they are singing together, or gradually, as a result of long-term training with the group, or both? In the present experiment, the subjects sang alone, so the results should reflect only the long-term training.

Subjects

In order to simplify the formant measurements, it was decided to use only bass singers, since their low fundamental frequencies give rise to spectra with closely spaced harmonics. Eight subjects participated, who had been singing in the same choir for several years (although not all of them at the same time). All were long-time residents of Stockholm, and had fairly similar Stockholm dialects in normal speech.

Method

The subjects were recorded one at a time. First, the subject spoke the text of a song phrase four times, and then sang it four times in accordance with the score using the correct key. The recordings were made in an anechoic room. The signals from a normal microphone and a

larynx wall accelerometer were recorded on separate tracks of a SONY F1 digital tape recorder (Fig. 1).

Fig. 1. Experimental setup.

A choral piece was chosen that was familiar to the subjects. The text contained several different sustained vowels in a fairly low register in the bass part (Fig. 2).

Fig. 2. Music score and phonetic transcription of the phrase used in the experiment. The nine vowels selected for analysis are marked with arrows.

Analysis

For each subject, wide-band spectrograms were made only of the third rendering of the spoken and sung phrases. The formant frequencies were measured manually by two experimenters from printouts of the spectrograms. The reading error was estimated to be less than 50 Hz.

For all selected vowels, the mean and standard deviation in the frequencies of formants 1-4 were calculated. Also, the difference $F_4 - F_3$ was calculated, offering a measure of the degree of clustering of these formants, and being thus closely related to the singer's formant.

Results

The reproducibility of the data was checked by comparing two versions of the same vowel, /o/, which occurred both near the beginning and near the end of the phrase. Table I shows the formant frequency averages across subjects for the first and second time when the vowel occurred during the subjects' reading and singing of the text. As can be seen in the table, the differences are merely small fractions of the standard deviations, so there were no significant differences in the formant frequency averages between the first and second occurrences (see Table I).

Table I. Formant frequency averages and standard deviations of the duplicate vowel /o/, first and second time it occurred in the texts. Population size = 8.

		SPEECH		SINGING	
		mean	SD	mean	SD
		(Hz)	(Hz)	(Hz)	(Hz)
F1	First	396	38	396	32
	Second	394	24	395	17
	Difference	-2		-1	
F2	First	741	57	767	52
	Second	731	74	774	44
	Difference	-10		7	
F3	First	2447	192	2349	91
	Second	2530	172	2406	149
	Difference	83		57	
F4	First	3178	68	2975	192
	Second	3108	209	3019	201
	Difference	-70		44	

Typical differences were found in the mean formant frequencies. Our averages and standard deviations for the first two formants of the entire set of vowels are plotted together with other previously published, comparable data in Fig. 3 in terms of formant charts. In our subjects, the means of the first and second formants were shifted in singing toward the center of the vowel triangle, i.e., the vowels were somewhat neutralized.

In a previous investigation, a similar comparison was made for four professional bass and baritone singers in a CVC context (Sundberg, 1970). Only the four vowels that occurred also in our experiment were

Fig. 3. Mean F1 and F2 (kHz) in speech and singing for (a) eight choir singer subjects, and (b) four professional singers as measured by Sundberg in 1970. Fig. 3c shows data from Rossing & al. (1986) comparing choir and solo modes of singing. In (a), the standard deviations in F1 and F2 are shown as ellipses. Arrows in the figure indicate the change from speech to singing.

shown. As can be seen in Fig. 3b, the results are similar. The second formant was lowered in the vowels /i/ and /ae/, while there were no clear changes in the vowels /o/ and /ø/. Thus, choral and solo singers seem to change their lower formants similarly when switching from speech to singing.

Fig. 3c, showing data from Rossing & al (1986), compares five singers' solo and choral singing. In this case, no systematic differences in F1 or F2 are apparent. We might therefore infer that, although speech and singing differ, singing is much the same as solo singing, as regards the first two formants.

Fig. 4 presents the corresponding data regarding F3 and F4. For the choir singers, it can be seen that the mean F3 was lower in singing, except for the vowel /ae/. Also, F4 was close to 3 kHz in all sung vowels, except /a/ and /ε/ which, interestingly, were the only two vowels that appeared in unstressed positions in the music. In professional singers (dash-dotted lines), both F3 and F4 seem to be fairly low already in speech; still, some reduction of F4 can be seen here as well.

Fig. 4. Vowel averages of F3 and F4 in speech (□) and singing (○). Each point represents the mean of eight subjects. For clarity, only the average standard deviations are shown (box). Data from four professional singers (Sundberg, 1970) are shown for comparison (dash-dotted lines).

While the two lowest formants define vowel quality almost entirely, the higher formants are more significant to the personal voice timbre. For this reason, averages of the higher formants were also calculated per subject (Fig. 5). In singing, all subjects but one lowered both their mean F4 and, to a lesser extent, also F3.

Fig. 5. Subject averages of F3 and F4 in speech (□) and singing (○). Each point represents the mean of nine vowels. For clarity, only the average standard deviations are shown (box).

An interesting aspect of formants number 3, 4, and 5 is the singer's formant which seems associated with a clustering of these formants (Sundberg, 1974). Therefore, the proximity of F3 to F5 or to F4 can be used as a pertinent measure. In many spectrograms, the fifth formant was hard to identify and measure and was therefore not included in the investigation. Fig. 6a shows a correlogram of the difference F4-F3 for speech and singing, for all vowels and subjects. There is only a weak tendency toward a smaller difference in singing. The corresponding data from Sundberg (1970) in Fig. 6b show a much clearer difference between speech and singing for professional singers. These results suggest that the choir singer subjects, unlike the professionals, have a similar level of their higher formants when they sing as when they speak. In other words, a prominent singer's formant appears to be the solo singer's characteristic.

EIGHT CHOIR SINGERS

FOUR PROFESSIONAL SOLO SINGERS

Fig. 6. Correlograms of the difference F4-F3 for all tokens; (a) results from choir singers in the present experiment, (b) data from professional solo singers (Sundberg, 1970).

The long-term-average spectrum (LTAS) level difference between the 3000 Hz and 500 Hz bands is a more direct indicator of the level of the singer's formant (Rossing & al., 1986). Therefore, LTAS were made of the entire material. The level of the 3 kHz band showed only a moderate relative increase (mean = 1.4 dB) in singing as compared to speech. This change could be ascribed to the average 7.2 dB increase of SPL observed in the sung case; it is well known that SPL increases are accompanied by a reduction of the overall spectrum tilt (Fant, 1959; Sundberg, Ternström, Perkins, & Gramming, 1986).

The question now is whether or not the subjects agreed more closely amongst each other on the formant frequency values in singing as compared to speech, i.e., if they tuned their formant frequencies to a pattern common to the group members. In order to investigate this, we adopted the standard deviations for the formant frequencies as a scatter measure allowing us to analyze the formant scatter across subjects. Fig. 7 shows a correlogram of the standard deviations across subjects for formants 1-4, comparing spoken and sung vowels. Since the standard deviations are shown in Hz, greater values are to be expected for the higher formants. Expressed in percent, the standard deviations were in the range of 5-10% for all four formants. It can be seen that the standard deviations of formants 1-3 were generally smaller in singing. Using a t-test for pair wise dependent measures, with the SD values as primary data, the SD of formants 1, 2, and 3 pooled was found to be significantly lower ($p < 0.01$) in singing. This supports the hypothesis that the subjects tuned their lower formants for the various vowels to patterns common for the group.

It is interesting that, according to Fig. 5, all subjects lowered their F4 in singing, while according to Fig. 7, the SD in F4 across subjects remained the same in singing. This suggests that, within subjects, F4 was changed systematically for singing, though to values that differed between subjects, such that the SD was unaffected when computed over subjects. Given the differences between individuals, it seemed pertinent to calculate also the formant scatter across vowels (Fig. 8). By and large, this scatter is higher than the scatter across subjects, since different vowels are being pooled with each other. We note that the SD's in F1 and F2, being mainly determined by the set of vowels used, are fairly independent of subjects. This explains the close grouping of the points for F1 and F2. Moreover, the SD values for F1 and F2 were slightly lower in singing also across vowels, cf., Fig. 3a. The SD in F3 and F4 are less vowel dependent and more personal, and show markedly lower scatter in the sung case. Since the scatter in F4 across subjects was unchanged in singing, it would seem that each choir singer subject has a systematic but personal way of putting F4 into a "singing" mode. To a lesser extent, this might hold also for F3.

Fig. 7. Standard deviations across eight subjects (log Hz) in F_1 , F_2 , F_3 , and F_4 in spoken and sung vowels.

Fig. 8. Standard deviations across nine vowels (log Hz) in F_1 , F_2 , F_3 , and F_4 for the eight subjects in speech and singing.

In summary, we have observed some clear formant frequency differences between our subjects' spoken and sung versions of the same vowels. As for F1 and F2, there was a tendency toward neutralization, and F3 and, particularly, F4 were both lowered in singing. The agreement between the subjects in F1, F2, and F3 was greater in singing. Within the subjects, the variability across the vowels of F3 and F4 was smaller in singing.

Discussion

With regard to the accuracy of the measurements, there were some differences between the spoken and sung vowels. In singing, the formant frequencies are more or less stationary for the duration of a tone, while in speech, the formants generally change quickly in frequency. Also, the syllable rate in speech is much higher than in the music chosen for this experiment. In speech, it was therefore often somewhat harder to judge the position of the formants from the spectrograms.

The higher speech rate would lead to a vowel reduction in some cases, i.e., neutralized vowel formant frequencies. However, our results exhibited a neutralization in the case of singing rather than in the speech, indicating that any effect of vowel reduction was overridden by the effect of changing from speech to singing.

It is interesting that the scatter across subjects in the three lowest formant frequencies was reduced in singing, even though the subjects did not perform in ensemble during the experiment. This suggests that singers habitually use particular modifications of the vowels when singing. As the fourth formant dropped in most subjects and vowels, it seems likely that there is a common articulatory background, such as a lowering of the larynx.

For the two lowest formants of most of the vowels studied, these modifications were similar to those made by professional singers. This evokes the question whether the demands raised in singing on the vocal apparatus impose articulatory restrictions entailing these vowel modifications. However, it is also possible that singers deliberately modify their vowels in singing in order to produce certain desirable timbral patterns. In any event, as the choral singers did not show any clear singer's formant, it can be concluded that they do not make these modifications in order to achieve a singer's formant.

The modifications of the vowels in singing reduced the formant frequency scatter across the subjects for the three lowest formants. Many choir leaders consider a uniform vowel articulation important. This could be for reasons of vowel intelligibility: agreement in formant frequencies would sharpen the spectrum peaks, thereby increasing the spectral contrast between different vowels and, possibly, facilitating vowel recognition. For vocal groups with one singer to each part, such

as barbershop quartets, similar spectrum envelopes should maximize beating in mistuned consonant intervals; in that case, the amplitudes of the common partials are more likely to match. This might give the singers better auditory cues to the correct intonation. In larger ensembles, on the other hand, beating must be abundant and probably too random to be of any use in aiding intonation. In cases where a formant is placed on a partial that is common to two simultaneously sounding tones, a small scatter may be helpful for intonation (Ternström & Sundberg, 1982).

The singer's formant is a voice property which usually is acquired by soloist training. Our choir singers did not show any prominent singer's formant. In a comparison of solo and choral singing, Rossing & al. (1986) found a much more prominent singer's formant even in the choral mode of singing. However, in that investigation, the subjects were singers who performed professionally both as solo and as choral singers, and the choirs they belonged to were of a high professional standard. This shows that the presence of a singer's formant differs between choirs. The present results would be more representative of the average amateur choir singer. If so, it would be inappropriate for a singer to sing with a singer's formant in an amateur choir. Doing so, this singer would fail to blend with the ensemble sound (cf., Sundberg, 1977). On the other hand, if all choir members sing with a singer's formant, which might occur in some opera choirs, the blending should not be affected. In any event, the degree to which a singer's formant is present in a choral sound must be relevant to the timbre of the choral sound.

The present investigation has demonstrated that choral bass singers for some reason adopt certain vowel formant patterns that are different from and more unified than those they use in speech. It remains to be investigated to what extent these unified formant patterns are shared also by the other parts of a choir. Also, it would be interesting to map the formant patterns characterizing different types of choirs. In such a study, it would certainly be rewarding to complement formant frequency measurement by synthesis and choral sounds.

Acknowledgments

The participation of the choir singers is gratefully acknowledged. Our thanks are also extended to the choral directors Margareta Skoglund and Lars Tillenius for their assistance. This project was supported by the Swedish Council for Research in the Humanities and Social Sciences.

References

Cleveland, T.F. (1976): "The acoustic properties of voice timbre types and their importance in the determination of voice classification", dissertation, University of Southern California, Los Angeles.

Goodwin, A. (1980): "Acoustic study of individual voices in choral blend", J. Res. in Singing 3:2, pp. 25-36.

Rossing, T.D., Sundberg, J., & Ternström, S. (1986): "Acoustic comparison of voice use in solo and choir singing", J.Acoust.Soc.Am. 79:6, pp. 1975-1981.

Rossing, T.D., Sundberg, J., & Ternström, S. (1987): "Acoustic comparison of soprano solo and choir singing", J.Acoust.Soc.Am. 82:3, pp. 830-836.

Sundberg, J. (1970): "Formant structure and articulation of spoken and sung vowels", Fol.Phoniat. 22, pp. 28-48.

Sundberg, J. (1974): "An articulatory interpretation of the singer's formant", J.Acoust.Soc.Am. 4, pp 838-844.

Sundberg, J. (1977): "The acoustics of the singing voice", Scientific American 236, March, pp. 82-91.

Sundberg, J., Ternström, S., Perkins, W.H., & Gramming, P. (1986): "Long-term average spectrum analysis of phonatory effects of noise and filtered auditory feedback", STL-QPSR 1/1987, pp. 57-80.

Ternström, S. & Sundberg, J. (1986): "Acoustics of choir singing" pp. 12-22 in S. Ternström (ed.) Acoustics for Choir and Orchestra, Publ. issued by the Royal Swedish Academy of Music, No. 52, Stockholm.

Ternström, S., Sundberg, J., & A. Colldén (1983): "Articulatory perturbation of pitch in singers deprived of auditory feedback" STL-QPSR 2-3/1983, pp. 144-155. (forthcoming in JSHR).