How loudly should you hear your colleagues and yourself?

Ternström, S. and Sundberg, J.

journal: STL-QPSR
volume: 24
number: 4
year: 1983
pages: 016-026

http://www.speech.kth.se/qpsr
II. MUSIC ACOUSTICS

A. HOW LOUDLY SHOULD YOU HEAR YOUR COLLEAGUES AND YOURSELF?
A study of SPL within choirs
Sten Ternström and Johan Sundberg

Abstract
A choir singer needs to hear both the sound of his/her own voice and the sound of the other choir members. If either of these signals becomes too loud, problems can be assumed to arise. The present study investigates the importance of the amplitude ratio between the sound of one’s own voice and the sound of an external reference signal.

Nine male choir singers were asked to sing in unison with synthetic reference vowel stimuli presented over ear-phones. In the ear-phones they also heard the sound of their own voice as picked up by a microphone in front of the mouth. The SPL of the reference stimuli was varied over a range of 40 dB, while the subjects were asked always to sing at a constant SPL as indicated by a level meter. The fundamental frequency agreement between the reference and the subjects’ responses was analyzed. The fundamental frequency of the responses depended strongly on the SPL of the reference in the case of the vowel /u/, presumably because of a pitch-amplitude dependence, which was compensated for. The fundamental frequency agreement between the reference vowels and the corrected responses was approximately constant over a dynamic range of 20 dB of the reference vowels.

Introduction

Choir singers are supposed to agree upon the pitch of the tones they sing. In order to achieve this agreement each choir singer is dependent on the perception of two signals. One is the auditory feedback signal, i.e., the sound which the singer perceives of his/her own voice. The other signal is the sound from all the companion singers in the choir. This latter signal can be regarded as an external reference, and the singer is required to bring his/her own voice pitch to agreement with the pitch of this reference.

The amplitude ratio between the external reference and the auditory feedback signal can be assumed to be relevant to the possibilities to attain pitch agreement in a choir. If the singer’s auditory feedback signal is much louder than the external reference, i.e., if the singer merely can hear his/her own voice, problems are likely to arise. Likewise, the situation would be problematic if the singer cannot hear his/her own voice because of a too loud external reference (Ward & Burns, 1978; Ternström & al., 1983). This situation can arise even though one’s own voice is always at an advantage because of both proximity and the private facility of bone conducted sound. Thus, the external reference has a double function. It serves as a reference but it also acts as a masker for the auditory feedback signal.

The amplitude of the external reference can be expected to vary considerably. Disregarding for the moment deliberate dynamic variations made for expressive purposes, the level of the external reference will
increase if the distance between the singers is decreased; also, it will increase in a reverberant room, and decrease in an environment poor in sound reflection, e.g., in an outdoor concert. Likewise, the amplitude of the auditory feedback will be influenced by the sound reflection in front of the singer.

The external reference acts as a masker for the auditory feedback signal, as mentioned. The most efficient masking effect is produced by the singer's section colleagues (e.g., your fellow tenors, if you are a tenor) because section colleagues normally sing the same sound and, also, they generally stand close. Choir leaders sometimes prefer to mix the choir sections, so that no singer is standing next to a section colleague. This should decrease the masking effect. Also it may affect the amplitude ratio between the auditory feedback signal and the external reference signal.

Thus, the amplitude ratio between the auditory feedback signal and the external reference must be important to the pitch agreement within a choir, and also it must be influenced by various factors. We may then ask how critical this ratio is: when do pitch matching problems start to occur and is there some optimal value of this ratio? The present investigation was carried out in order to elucidate the effect of this ratio on pitch agreement within a choir.

The main idea was to study the pitch behavior of single choir singers under conditions of a systematically varied amplitude ratio between the auditory feedback and the external reference. The subjects task was to sing in unison with synthetic reference vowels which were presented over ear phones at differing, though realistic, sound pressure levels. The subjects were required to sing at a given SPL regardless of the stimulus level. In order to present the external reference tones at realistic levels, it was necessary to know what the typical SPL-values are within a choir. This was measured in a special experiment.

**SPL-values of external reference**

An SPL-calibrated stereophonic tape recording was made during two choir rehearsals by means of a pair of binaural headborne microphones (Sennheiser MKE2002). A total of about 30 minutes of effective rehearsal time was recorded in various positions within each of two choirs. One choir was a highly regarded amateur choir rehearsing in a reverberant hall, and the other was a professional choir rehearsing on the stage of a concert hall.

The measurements showed that a mezzoforte typically produces an SPL of 90-95 dB at 15 cm from a bass singer's mouth. This corresponds to an SPL of about 80 dB at the ears of the choir singers. The highest values were observed in the soprano section, where the SPL occasionally reached 115 dB on high notes. The SPL data were evaluated in terms of histograms. The histograms for the various positions did not appear to differ substantially with respect to modes, averages and extremes. The histogram in Fig. 1 pertains to the pooled data for the various positions.
within each of the two choirs. It can be seen that both choirs yielded 
an average and a mode value close to 80 dB; the standard deviation is 
about 10 dB, and the extreme values are close to 60 and 100 dB.

![Histogram of SPL values](image)

**Fig. 1.** Histograms of SPL values recorded in different positions within 
each of two rehearsing choirs (solid and dashed lines, respectively).

**Method**

The audibility of one's own voice over a choir must depend not only 
on the SPL difference, but also on spectral characteristics. The sound 
of one's own voice dominates at low frequencies because of the lowpass 
nature of bone conduction and because of the directivity of the sound 
radiated from the mouth at high frequencies. The preferred level of 
the external reference is, thus, likely to depend on the vowel spectrum. 
For this reason we chose two stimulus vowels: /u/, which is poor in 
spectral components at high frequencies and /a/, which is more wealthy 
in high frequency spectral components.

A test tape was prepared including these two stimulus vowels which 
were synthesized on a minicomputer, see Fig. 2. The fundamental frequen-
cy was in the vicinity of the pitch of G4 (196 Hz), and a vibrato having 
an extent of ± 1% was added in order to eliminate the possibility that 
the subjects used beats to tune their fundamental frequency. The tape 
included 22 stimulus tones, 11 /u/ and 11 /a/, all of 9 sec duration. 
The fundamental frequency differed slightly between adjacent stimulus 
vowels on the tape so that the subjects could never use exactly the same 
pitch in two neighboring stimuli on the tape. The fundamental frequen-
cies were chosen within a range of 240 cent.

As the external reference in reality varies within a rather wide 
dynamic range (vide supra), an SPL range of at least 40 dB was judged 
desirable for the stimulus tones. Since no tape recorder available 
could meet this demand retaining an acceptable signal-to-noise ratio, a 
special variable-amplification device was constructed. The stimuli were 
all recorded at maximum amplitude on the tape and then played to the 
subjects via a voltage-controlled amplifier. The control voltage for
Fig. 2. Examples of spectra of the two stimulus vowels used in the test.
this amplifier was taken from a frequency-to-voltage converter, which in turn was fed by a sinusoidal control signal recorded on another track of the same tape. The accuracy obtained was better than 1 dB. A total of 9 SPL values were chosen within the range of 40 dB. In the stimulus tape two SPL values were duplicated for both vowels.

A block diagram of the experimental setup is shown in Fig. 3. The experiment was carried out in an anechoic room. The subjects were seated in front of a microphone, a level meter, and a small lamp which was lit as long as a stimulus tone was sounding. The level meter carried a fixed scale mark which indicated the SPL to be produced by the subject for all external reference vowels, regardless of their loudness. This SPL value was chosen to 90 dB at .15 m from the mouth. The subjects heard these stimulus vowels over the ear-phones. In the ear-phones they also heard their own voice as picked up by the microphone in front of the mouth. In this simulated auditory feedback loop, the gain was regulated so as to give a realistic level of the auditory feedback in an ordinary room. Ideally, the normal acoustic mouth-to-ear transfer should have been modelled here. However, this proved to be complicated, as this transfer varies with vowel (cf., Lindqvist-Gaufin & Sundberg, 1974). Thus, for the sake of simplicity this signal was merely lowpass filtered at 3.4 kHz. The spectra of the external reference vowels were lowpass filtered at 3.8 kHz.

The subjects were instructed

- to sing in unison with the stimulus tone they heard in the headphones, using the same vowel as that of the stimulus tone;
- to maintain the SPL indicated by the mark on the level meter irrespective of the loudness of the sound in the headphones;
- to keep a constant distance of .15 m to the microphone;
- to sustain the tone as long as the lamp remained lit.

Each subject had a short practice run with three or four stimuli before the actual test began.

The tones sung by the subjects and the reference tone that they heard in the headphones were both recorded on two tracks of a tape, so that errors in the speed of the stimulus tape could be compensated for.

The experiment was run with nine subjects, who all were highly experienced choir singers.

Results

The pitch behavior in all responses was examined by means of a hardware pitch extraction combined with a computer program which statistically measures and treats fundamental frequency data (Ternström, & al., 1983). The vibrato, if any, was eliminated by means of a suitably defined averaging filter. The linear average and the standard deviation of fundamental frequency in each response was determined from
Fig. 3. Block diagram of the experimental setup.
fundamental frequency histograms. The average fundamental frequency of the reference tone stimuli were determined in the same way.

Thus, the statistical analysis of the responses yielded two measures, the average and the standard deviation of the fundamental frequency of the 9 sec long responses. The standard deviation is strongly dependent on occasional fundamental frequency excursions, e.g., in the onset or the decay of the tone. Also, it is dependent on whether or not the vibrato, if any, was completely eliminated. For these reasons the average of the fundamental frequency must be considered as a more relevant measure.

In many instances the fundamental frequency of the response showed a strong correlation with the stimulus SPL. This effect was more frequent for the vowel /u/ than for the /a/. The general trend was that the average fundamental frequency produced in the response vowel dropped for rising stimulus SPL. This was interpreted as a pitch-amplitude effect which has been observed previously both for sinewaves (Ward, 1954) and complex tones (Sundberg & Lindqvist-Gauffin, 1973). This effect was compensated for by computing the linear regression coefficient for each subject's productions in each vowel and then correcting the measured fundamental frequency value by the value predicted by the best linear fit line. This correction was applied in all cases where there was a statistically significant correlation between stimulus SPL and response fundamental frequency.

After this, the subject's average fundamental frequency was compared with the average fundamental frequency of the stimulus as recorded on the tape, and the difference between these averages was computed and expressed in per cent. This percentage thus shows the fundamental frequency discrepancy between the external reference and the subject's response. It can be assumed to reflect the difficulty which the subject experienced in producing the same fundamental frequency as the external reference stimulus. Also, the absolute value of this frequency discrepancy would be related to the pitch agreement within a choir section.

Fig. 4 shows the standard deviations and the averages, computed over the nine subjects, of the difference between the averaged fundamental frequency of the reference tone and the averaged fundamental frequency of the subject's response. The values are given in absolute numbers. The standard deviations are substantial, suggesting that the subjects reacted very differently to the various reference tone conditions. Thus, while some subjects' fundamental frequency averages did not deviate at all from the mean fundamental frequency of the reference, other subjects displayed more or less large errors. It is important here to recall that for most SPL values the subjects had a reduced possibility either to hear the sound of their own voices or the sound of the reference signal. Under such conditions it is likely that the pitch behavior shows a great intersubject variability, probably reflecting various personal factors of significance to phonatory pitch control. Thus, it seems fair to conclude that the great standard deviations are
explicable and that they do not reduce the interest of the averages.

From Fig. 4, it can be seen that for both vowels the two duplicate stimuli yielded similar response data; the agreement is within one or two percent. This supports the assumption that these averages are rather reliable.

It can also be observed that the results are rather similar for both vowels. Both curves show a slow increase of the frequency error towards softer external references and a steep increase of the same error for the loudest external reference. Inbetween these values the errors remain close to about .5% within a range from 75 to 95 dB SPL. This supports the assumption that the loudness of the external reference may vary from -15 dB up to +5 dB relative to the SPL produced by the singer without causing serious problems for the singer to bring his fundamental frequency to agreement with the rest of the choir.

However, there is some slight differences between responses collected for the two vowels which seem worthwhile to observe. In the case of the /u/, the subjects’ agreement with the external reference is slightly better for the loudest reference SPL than in the case of the vowel /a/. This probably reflects the fact that the vocal effort needed to sing at a given SPL is greater in the vowel /u/ than in the vowel /a/, because of the lower first formant frequency in the /u/ (see Fant, 1960). Thus, when the subjects sang the /u/, they had probably more bone conduction to help them hear their own voice, even when the reference was loud. On the other hand, when the reference was soft, the /a/ should have been causing less problems than the /u/, but this does not seem to be the case. On the contrary, the vowel /a/ seems to have been the harder vowel to match for soft references, even though the differences between the two curves are small.

Discussion and Conclusion

Some (but not all) subjects’ fundamental frequency response showed a marked dependence of the amplitude of the external reference. This effect varied greatly between subjects and seemed even to be absent in some subjects. This effect may be relevant in choir singing. It suggests that the singers will respond with a fundamental frequency which disagrees more or less with the fundamental frequency of the signal they hear from the rest of the choir. This would lead to a tendency for the entire choir to change its reference pitch during the performance of a piece. The effect should then be possible to reduce by changing the amplitude of the individual singer’s external reference, i.e., by spacing the singers more, by bringing them closer together, or by changing the acoustic properties of the room. The fact that the effect varies greatly between different subjects suggests that the total effect would depend on how influential those choir members are (e.g., because of vocal loudness) which possess a marked pitch-amplitude dependence. Thus, the effect may depend strongly on one or few singers.

Although care was taken to achieve realistic conditions, the exper-
iment was artificial in certain respects.

The reference tones were synthetic, and, hence, lacked the waveform perturbation and other irregularities, that characterize natural vowel sounds. Some of the disadvantages of this regularity was eliminated by the vibrato. The regularity of the vibrato, on the other hand, was also unnaturally high and this might have offered a pitch agreement clue to the singers, which does not exist in reality. Moreover, the external reference was represented by one single tone in the present experiment, while in reality it is produced by several singers, and hence, probably less well-defined. Thus, in our experiment the reference was more regular than in reality. Presumably, this simplified the task to match the pitch of the reference somewhat. Also, in our experiment the singers could devote all their attention to the task of matching the pitch of the external reference. In real choir singing, there are other tasks that require the singer’s attention as well, e.g., reading the music, observing the signs from the leader etc. All this leads us to assume that in a real choir situation the smallest average fundamental frequency error is probably greater than the .5% that we found in our experiment.

The tolerance range for the external reference was found to be about -15/45 dB relative to the SPL that the singer’s own voice generated at the singer’s own ear. In reality, this last mentioned SPL value is not well defined. It will be influenced by the presence of sound reflectors near the singer. In our experiment it was adjusted subjectively. If the auditory feedback level is raised, the pitch matching difficulties will occur at higher levels of the external reference, and vice versa. This implies a transposition of the curves in Fig. 4 to the right or to the left.

The above considerations lead us to suspect that, in reality, the fundamental frequency disagreement within a choir is probably greater than what we have measured. Also, the curves in Fig. 4b might be shifted to the right or to the left, and they probably differ slightly between different vowels. On the other hand, their U-shaped form with a flat minimum of about 20 dB width can be regarded as typical; hence, we conclude that the amplitude ratio between the external reference and the auditory feedback can be varied within a 20 dB range without causing serious problems regarding pitch agreement within a choir. In other words, this amplitude ratio is not very critical.

The present study is merely a first attempt to gain an insight into feedback and reference signals in choir singing, and several other experiments are needed. How much can the amplitude ratio between the external reference and the auditory feedback signal be varied in practice? Would the curves in Fig. 4b change if the singers were to sing at a louder SPL? To what extent is the average pitch matching error representative of other aspects of acoustic agreement within a choir, e.g., synchronization and timing? All these questions are important but at present we must leave them to future research.
Acknowledgments

The present work was supported by The Swedish Council for Research in the Humanities and Social Sciences. The participation of the choir singers in the experiments is gratefully acknowledged.

References


