AUDITORY AND KINESTHETIC FEEDBACK IN SINGING – SIGNIFICANCE AND EFFECTS OF TRAINING ON PITCH CONTROL*

D. Mürbe¹, ², G. Hofmann¹, ², F. Pabst², J. Sundberg³

¹Department of Otorhinolaryngology, Technical University of Dresden, Dresden, Germany
²Voice Research Laboratory, University of Music Carl Maria von Weber, Dresden, Germany
³Department of Speech, Music and Hearing, KTH, Stockholm, Sweden

* Parts of this investigation were presented at the 31th Annual Symposium Care of the Professional Voice, Philadelphia June 2002 and are accepted for publication in the Journal of Voice

Abstract: An accurate control of fundamental frequency is one of the essential demands in professional singing. This control relies on auditory and kinesthetic feedback. However, a loud accompaniment may mask the auditory feedback, leaving the singers to rely on kinesthetic feedback. The object of the present study was to estimate the significance of auditory and kinesthetic feedback to pitch control in 28 students beginning a professional solo singer education. Since it seems reasonable to assume that pitch control can be improved by training, the same students were reinvestigated after 3 years of professional singing education. In both parts of the study the singers sang an ascending and descending triad pattern with and without masking noise in legato and staccato and in a slow and a fast tempo. Fundamental frequency and interval sizes between adjacent tones were determined and compared to their equivalents in the equally tempered tuning. The average deviations from these values were used as estimates of intonation accuracy. For both parts of the study, intonation accuracy was reduced by masking noise, by staccato as opposed to legato singing and by fast as opposed to slow performance. After education, the contribution of the auditory feedback to pitch control was not significantly improved while the kinesthetic feedback circuit was improved in slow legato and slow staccato tasks. The results support the assumption that the kinesthetic feedback contributes substantially to intonation accuracy and might be improved by training.

Keywords: singing, pitch control, training, auditory feedback, kinesthetic feedback

I. INTRODUCTION

The high demands on intonation in professional singing require precisely acting pitch control systems. Auditory and kinesthetic feedback of the phonatory system have been described to contribute to singers’ pitch control [1, 2].

Auditory cues are commonly regarded as the obvious tool for pitch control in singing under normal circumstances. However, auditory feedback cannot explain the fact that singers are able to continue phonating accurately even when they cannot hear their own voices. This situation is typically experienced in solo singing when the orchestral accompaniment is loud; SPL values as high as 110 dB have been observed on orchestral podiums [3]. Under such conditions, singers have to rely on the performance of a second intraphonatory feedback circuit, based on kinesthetic discharges.

The aim of the present study was to estimate the importance of auditory and kinesthetic feedback to pitch control in students beginning their professional solo singer education. The effect on pitch control was investigated in tasks differing in complexity, such as legato or staccato, or slow and fast singing.

The effects of a professional training of the singing voice should include a sufficient accuracy of intonation. A longitudinal approach, in which the singer is used as his/her own control would represent a promising opportunity to test the effects of training. Therefore, the singing students were reinvestigated after 3 years of education to assess the effect of training on pitch control in singing.

II. METHODOLOGY

In the initial investigation 28 singing students were examined at the beginning of their professional solo singer education at the University of Music Carl Maria von Weber, Dresden [4]. After 3 years of professional solo singer education, 22 students, 13 female and 9 male students, mean age 24,0 ± 1, 6 years, still continued their studies and could be re-investigated [5].

Subjects were asked to sing an ascending and descending triad pattern up to the twelfth and back on the vowel [a:] at a moderate degree of vocal loudness. The starting pitch, chosen so as to fit comfortably the pitch range of the individual subject, was given by means of a synthesizer. Each subject sang the sequence twice, first without masking noise, and immediately...
afterwards with a masking noise presented via headphones. The masker was a white noise band-pass filtered (24 dB/octave) at 50 Hz and 2000 Hz. The SPL of the noise was 105 dB. The masking efficiently eliminated the auditory feedback.

The sequences without and with masking noise were recorded in different conditions: a) legato slow, b) legato fast, c) staccato slow, d) staccato fast. The slow and fast tempi corresponded to metronome settings of 40 and 160 beats per minute, respectively. The output from a portable electroglottograph (EGG) (Laryngograph, London, UK), and the audio signal as picked up by a microphone (distance to mouth 0.3 m) (ECM-959DT SONY, Japan) were recorded on a digital audio tape (TCD-D10, SONY, Japan). The identical test program was recorded again after training.

Figure 1. F0 contour of a recorded sequence.

Fundamental frequency (F0) was mostly estimated from the EGG signal using the Soundswell workstation program package which also displayed the resulting F0 contour on the computer screen (Fig.1) (Soundswell, Solna, Sweden) [6]. In some of the female subjects the EGG signal produced errors in the F0 measurement at high pitches. In such cases F0 was measured from the audio signal. For determining the mean F0 for each pitch, a set of complete vibrato cycles was selected from the quasi-steady state section, thus excluding onset and offset transients. The frequency distribution of this selection was analyzed, using the histogram module in the Soundswell package, which also displays the mean F0. The mean F0 of each tone was measured.

The sizes of the 10 intervals included in each triad sequence were determined by calculation of the F0 interval between adjacent tones, expressed in the logarithmic cent unit. The absolute values of the deviations of these intervals from their equivalents in the equally tempered tuning, henceforth the interval deviations, were determined and regarded as a measure of the accuracy of intonation. The averaged interval deviation of the 10 intervals contained in a complete triad sequence was defined as the mean interval deviation.

Interval deviation data were referred to a statistical analysis carried out by means of a repeated measures design (ANOVA), with time (before/after), masking (without/with masking), technique (legato/staccato) and tempo (slow/fast) as within subject factors.

III. Results

The measurements before the professional singing education showed a significant difference between the unmasked and masked conditions ($p<0.001$), mean interval deviations across all subjects amounting to 33.3 and 47.3 cent, respectively. The effect of masking appeared to be independent of technique and tempo. Figure 2 illustrates these results for the different conditions in terms of the distribution of individual mean interval deviations. Further, a significant difference was found between legato and staccato performances ($p<0.001$) as well as between slow and fast performances ($p<0.001$) [4].

Comparison of the before and after education measurements did not show a general difference between these conditions. For the after education measurements, the masking increased the mean interval deviation across
all subjects from 35.3 cent to 45.1 cent [5]. Statistically, this effect of masking on pitch accuracy did not differ significantly between the before and after education measurements ($p = 0.15$). However, according to the ANOVA, there was a significant interaction effect of “time” and “tempo” ($p = 0.001$), reflecting different effects of education for the slow and fast performances. Intonation accuracy improved for the slow performances, the mean interval deviation across all subjects dropping from 37.7 cent before education to 32.7 cent after education. Fig. 3 shows the distribution of individual mean interval deviations for all slow performances, before and after education. The strongest effects appear for the masked test conditions, both for legato and staccato performances. No improvement of intonation accuracy was found for the fast performances after education.

Figure 3. Comparison between before and after education data in terms of a box plot diagram showing the distribution of mean interval deviations (cent) for all slow tempo data (subjects n=22).

IV. DISCUSSION

The present study was carried out to assess the significance of auditory and kinesthetic feedback on pitch control in singing and to investigate effects of training on both feedback circuits. The slow and fast as well as the legato and staccato conditions were included in our experimental design since they raise different demands on pitch control.

intonation accuracy was found to be reduced by masking noise, by staccato as opposed to legato singing and by fast as opposed to slow performance. The masked and unmasked conditions allow an insight regarding the roles of the auditory and kinesthetic feedback systems in pitch control. Auditory feedback is commonly regarded as the main tool for pitch control in singing [7, 8]. However, under certain circumstances singers cannot hear their own voices, because the auditory feedback temporarily might be masked by the choral sound of the fellow singers or a loud orchestral accompaniment [3, 9]. A significant effect of masking was observed, amounting to a mean deterioration of pitch accuracy by 14 cent at the beginning of the students’ professional solo singer education [4]. This effect was only slightly smaller (10 cent) after education, a statistically non-significant difference. This suggests that the auditory feedback contributed to pitch control to a similar degree before and after education. The effect of masking was similar for the various tempo and technique conditions, see Figure 2. Therefore, the differences in intonation accuracy associated with these conditions should reflect the importance of the kinesthetic feedback.

The kinesthetic feedback circuit, a complex neuromuscular reflex system, depends on discharges of mechanoreceptors, mainly located in the intrinsic laryngeal muscles, the subglottic mucosa and the laryngeal joints [10, 11]. The afferent discharges from these receptors are fed back to the motoneurone pools in the brain stem operating as individual controllers for laryngeal action and to the overriding subcortical system [1]. Within the masked condition, intonation accuracy differed between the various tempo and technique conditions; a greater mean interval deviation was observed for the staccato than for the legato condition and also for the fast as compared to the slow conditions (see Fig.2). In a staccato performance singers would need to rely on an absolute neuromuscular memory of pitch while in a legato performance they could recruit also a relative neuromuscular memory [12]. The difference observed between staccato and legato performances suggests that the former memory is less precise than the latter.

Comparing data recorded before and after education, a significant improvement of pitch accuracy was found after education for the slow performances. For instance, for the masked slow staccato condition a mean pitch accuracy improvement of 9 cent was found after education. For the same condition, a study carried out by Ward and Burns showed a 17 cent better pitch accuracy in singers than in untrained subjects [2]. The difference between their results and our findings appear expected, given the fact that they compared singers and nonsingers. The improvement of intonation accuracy observed for the masked slow staccato task indicates that the accuracy of the absolute neuromuscular memory of pitch increased after education. Incidentally, this ‘absolute kinesthesia’ is important not only to staccato performances, where adjacent tones are separated by a pause. It is also essential for intonation at the beginning of a phrase, if no rehearsal of target pitch is allowed. In fast singing our study showed no improvement or even a modest impairment
was observed. Probably, a period of 3 years of professional training might not be long enough to improve pitch control in demanding vocal tasks such as fast singing. Also, the accuracy of measurement is smaller for short than for long tones; the shorter the tone sequence, the more difficult the pitch extraction.

It is interesting that our study showed no training effect for the basic, most easy condition – the unmasked slow legato. This task – singing slowly a triad or scale with normal auditory feedback – may reflect the limit of intonation accuracy, which would be reached early in any singing education. Finally, it is worthwhile to emphasize that, on average, the intonation errors were only slightly (10 cent) greater when the auditory feedback was eliminated. This implies that the kinesthetic feedback, contributes substantially to intonation accuracy.

V. CONCLUSION

The present investigation has shown that singers’ intonation accuracy is reduced in the absence of auditory feedback. Under such conditions, the singers have to rely on kinesthetic feedback circuits. The performance of this feedback is significantly affected by the task that the singer performs. Thus, the mean intonation error was greater in fast than in slow singing. It was also greater in staccato than in legato singing. Professional solo singer education did not significantly affect the contribution of the auditory feedback to pitch control in singing. Such education seems mainly to affect intonation accuracy in terms of an improved accuracy of the kinesthetic feedback circuit.

REFERENCES