MOTOR REPRESENTATIONS AND THE EFFECTS OF AUDITORY FEEDBACK DISRUPTION ON SINGING REMEMBERED TUNES

By

Aysu Erdemir

Thesis

Submitted to the Faculty of the

Graduate School of Vanderbilt University

in partial fulfillment of the requirements

for the degree of

MASTER OF SCIENCE

in

Psychology

August, 2010

Nashville, Tennessee

Approved:

Professor John Rieser

Professor Daniel Ashmead

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER I

INTRODUCTION

This study is about how people know how to sing remembered pitches. In part, people have to know what actions to produce in order to sing remembered tunes, and in part they need to rely on hearing themselves when they sing. Thus, the current study seeks to examine the importance of motor and auditory representations for singing and pitch memory by blocking-out auditory feedback mechanism.

The human voice is capable of producing a vast array of sounds ranging from speech sounds to more musical ones; and the first musical instrument is believed to be the human voice itself. Also, singing is different from other types of music performance in regards to how the sound is produced; the source of the sound is within the body over which the singer has to have precise control. Moreover, singing, unlike any other instrumental music, appears to be universal - practiced across all known cultures. Despite the commonality and importance of singing however, it is surprising to see how little is known about the basic processes underlying this complex action.

In principle, successfully executing a motor action depends on efference, the motor commands from the central nervous system (CNS) to the periphery, and the afference, motor/proprioceptive (sensory) feedback from the muscles/joints toward the CNS. However, the successful completion of a musical action, such as singing, is complicated by the involvement of the auditory modality, specifically auditory feedback and memory representations for the melody. Thus, producing a succession of sounds with the voice is a multifaceted motor action which requires simultaneous and coordinated use of bodily structures, an efficient use of auditory and motor/proprioceptive feedback as well as a solid memory representation for the spectral–temporal characteristics of the melody.

The core skill required for singing is the control of the fundamental frequency (F0) for the voice. Hence, motor representations & feedback act together with auditory feedback to stabilize voice fundamental frequency (F0) during vocal production (Figure 1). Most research, however, on vocal production until this point has focused on the role of auditory feedback, while undermining the potentially significant influence of motor representations and feedback. Hence, the current study is specifically designed to reveal the importance of motor representations and motor feedback for vocal performance and memory under conditions of masked auditory feedback.

Figure 1. Representation of Auditory and Proprioceptive Feedback Mechanisms Involved in Singing. Auditory feedback is both air and bone conducted, whereas proprioceptive feedback involves feedback from the joint & muscle receptors.

Auditory Feedback and its Importance for Vocal Performance

Auditory feedback includes the audible consequences of an action; simply put, it is hearing someone's own vocal production. However, unlike the acoustic signal that is heard from outside sources, the auditory feedback is both air and bond conducted. Auditory feedback is received through the outer ear and through the bony structures of the skull that emphasize low frequencies. It is the kind of feedback directly available to people's awareness when producing sounds.

Many clinical and experimental studies have shown auditory feedback to be essential for maintaining normal voice control in general; and most of these studies focus on the cases of hearing loss to reveal the importance of auditory feedback. Hearing loss that occurs early in life, for instance, causes serious distortion for speech and it hampers normal language development. There is an observable difference between deaf infants and hearing infants babbling (Oller and Eilers, 1988), and deaf infants display distorted syllable production in the first year of life, indicating that audition plays an important role in normal vocal and language development. Even postlingually profoundly deaf adults have abnormal levels of F0, variations in F0, difficulties controlling vocal intensity and speaking rate (Leder et al., 1987; Cowie and Douglas-Cowie, 1992); and some sounds, notably /s/, deteriorate after a period of deafness (Raphael et al., 2007). Other studies have demonstrated that masking of auditory feedback under controlled laboratory studies leads to deterioration in intonation accuracy, pitch-matching accuracy and fine control of F0 (Elliot and Niemoeller,1970; Mürbe et al., 2002). Moreover the studies in which auditory feedback is altered systematically provide further support to the idea that people normally rely on pitch feedback to control voice F0. For instance, presenting sudden alterations in auditory pitch feedback while subjects are producing a sustained sound, cause compensatory changes in the voice F0 (Burnett et al., 1998; Jones and Munhall, 2002; Jones and Keough, 2008; Xu et al., 2004).

Although, many studies have documented a clear role for auditory feedback to control F0, it is also argued that the auditory system may not serve as an effective feedback mechanism for 'skilled' articulation since the information it provides arrives CNS late (Raphael et al., 2007). For instance, auditory feedback cannot explain how professional singers are able to preserve their intonation accuracy even when they cannot hear their own voices accurately. This situation is naturally experienced by professional singers on stage performances, when the auditory feedback is naturally masked from the fellow singers or by a loud orchestral accompaniment.

Motor Representations and Proprioceptive Feedback, their Importance for Vocal Performance

Singing requires a high degree of muscle coordination, so motor/proprioceptive information also plays an important role in vocal control. The motor information used during singing includes the intentions and motor plans that control the muscles and cause them to contract as well as the feedback from the joint & muscle receptors (proprioceptive feedback) which provides information with respect to muscle movement, position and length changes. The proprioceptive feedback involved in singing provides feedback with regard to the relative positioning of articulators such as larynx, teeth, tongue, hard/soft palate, and other organs of speech. It allows us to sense the velocity and direction of movement, and positioning of bodily structures.

Proprioceptive pitch control includes intrinsic muscles of the *tongue* and *facial muscles* which provide useful information for muscle length changes; *laryngeal* muscles that control the elasticity and tension of vocal folds, so that we can make changes to the frequency of the sound we produce; and finally the *respiratory muscles* that help to move the chest wall so that we can regulate breath pressure (Raphael et al., 2007).

This type of feedback (direct feedback from the muscles) is delivered more quickly than external auditory feedback, and unlike auditory feedback, it is less available to conscious awareness; yet it is very important for speech and vocal performance in general. Sensation of movement and positional feedback is highly necessary for the fine control needed in skilled motor actions, such as singing.

Unlike non-singers, trained singers are highly aware of the muscular movements because teachers of singing play very close attention to what their students "feel" within their body. Classical vocal training focuses on teaching singing through sensations, that is each pitch is felt at a distinct and different place in the body (Di Carlo, 1994) and singers should be constantly thinking about the kind of sensations they are feeling while they are singing (Appelman et al., 1986). In this type of training the whole body is represented by a road map where each pitch has a specific sensation attached to it; and the students are expected to memorize the sensation attached to a particular pitch; so they learn to recognize and distinguish pitches not by how they sound to them, but by how they feel internally – an ability which requires extensive vocal practice. Hence, a higher reliance on auditory information seems to be a characteristic of less experienced singers, while more experienced voices rely heavily on their use of proprioceptive feedback and motor memory of 'pitch placement' in the vocal tract.

Direct support for this idea comes from the comments of a professional opera singer, the only deaf mezzo soprano in the world, Janine Roebuck, who was classed as severely deaf at age 28. She reports how she relies on "feeling" of notes to accurately produce sounds:

> *"I stopped listening to myself, which is bad technique in any case, because everyone has a skewed perception of how their voice sounds. My singing teacher taught me to trust my technique, sing on the breath and from the heart."* (Roebuck, 2007).

Di Carlo (1994) argues that, unlike in speech, auditory feedback does not provide a reliable means for voice control for singers. The singers have skewed perception of what they are hearing, because the sounds return to the singers' ears after having been modified by the acoustic characteristics of the surroundings. The surroundings mainly transform the timbre of the sound perceived by the singer (by causing some of the harmonics to be reinforced or attenuated), yet timbre contains important acoustic cues indicating the placement of the voice (Di Carlo, 1994). So professional singers have to learn that they cannot rely on auditory feedback heavily in cases where room acoustics is different.

Moreover, performing in a choir or with an orchestra requires constant monitoring of proprioceptive feedback (Murbe et al. 2002). In many situations such as when performing in a group of other singers or with a loud orchestral accompaniment, singers are often left unable to hear their own voices due to the natural masking from other singers or a loud orchestral accompaniment. In those situations, where auditory feedback is masked naturally, singers have to rely solely upon an internal motor feedback circuit to control voice. Because of their ability to use motor information, experienced singers manage to sing in tune even when the orchestra or the choir is loud.

In line with the ideas presented above, Murbe et al. (2004) has shown that after three years of professional training, singers have improved performance under conditions of masked auditory feedback. It has been suggested that the extensive vocal practice and professional training singers receive improves their ability to use motor memory and proprioceptive feedback to control their pitch in situations where auditory feedback is limited. The study also suggested that intonation accuracy under conditions of masked auditory feedback is subject to the level of training of the singer, with more training yielding better ability to control pitch without auditory feedback.

Absolute versus Relative Memory for Pitch and Musical Time

The pitch of a melody can be encoded and remembered in two different ways, either in absolute or in relative terms. An *absolute pitch* code consists of the fundamental frequencies (F0) of each individual tone in a melody. A *relative pitch* code, on the other hand, does not contain information about the individual fundamental frequencies; rather it consists of intervals – pitch distances between successive tones in a melody – thus, relative pitch information is independent of the key in which the melody is sung.

This distinction between relative and absolute pitch is apparent in the fact that people have no trouble recognizing songs in transposition (Deutsch, 1978; Levitin, 1994). For example, a melody such as Happy Birthday can be transposed to different keys, but as long as the relative pitch distances between tones are held constant, it is recognized as the same melody.

Encoding musical memory in two different ways has also been demonstrated for the temporal features of music. Similar to pitch codes, the time in music can also be encoded and remembered in either absolute or relative terms. The two main elements that describe musical timing are tempo and rhythm: *Tempo* refers to the overall timing of music, the rate at which people would tap their foot to it, how quickly or slowly it goes by. *Rhythm*, on the other hand, is the movement marked by the relationship between the durations of successive notes. Thus, musical tempo is most naturally described in absolute terms, whereas rhythm is most naturally described in relative terms (Monahan, 1993).

Hence, similar to transposition in pitch, the overall timing – musical tempo - of a melody can also be easily transformed, but it is still recognized as the same melody as long as the relation between rhythmic elements (rhythmic pattern) is held constant (Monahan, 1993; Serafine, 1979). Thus, the identity and recognizability of a melody is maintained through transposition in pitch and changes in tempo (Monahan, 1993).

The question of whether we encode musical pitch and time in relative or absolute terms is discussed often. As we attend to a melody do we register the absolute information, such as pitch and tempo, or do we calculate melodic intervals and tempo-free rhythmic information? There is considerable evidence in favor of the idea that both infants and adults encode and store melodic information primarily in terms of relative values (Plantinga and Trainor, 2005), as also indicated by the fact that people can readily recognize a melody in transposition and played at different tempo. People has shown to exhibit poor absolute memory for isolated pitches (Deutsch et al., 1999) and event-related potential (ERP) responses to pitch changes have also shown that most people primarily use a relative code to encode pitch (Trainor et al., 2002). However, on the other hand, people also have been shown to possess good absolute memory for the pitch and tempo of highly familiar songs that are heard and played always in the same key and at the same tempo (Levitin, 1994; Levitin and Cook, 1996; Schellenberg and Trehub, 2003). So the experience of a particular musical melody, always at the same pitch and at the same tempo, seems to be leading to some absolute pitch and time retention.

In sum, although music, for the most part, places significant emphasis on relative properties (pitch relations and rhythmic patterns) which essentially form the basis of a musical structure; if absolute information is reinforced through extensive exposure, then people exhibit surprisingly good absolute memory. This distinction between absolute versus relative pitch and time encoding is also important for the purposes of this study, as is for pitch and time memory in general. Although aforementioned studies suggest a distinction for absolute and relative pitch, these studies do not uncover the nature of memory representations, that is, whether absolute or relative pitch and time representations are heavily dependent on auditory or motor modalities.

The Purpose of the Study

The main purpose of this study is to investigate the role played by auditory and motor systems in the skillful control of singing for trained-singers, instrumentalists and people with little or no musical training, when auditory feedback is masked in different ways. Specifically the study focuses on how absence of auditory feedback influences the accuracy of singing in terms of *relative & absolute pitch*, *tempo* and *rhythm*; and whether or not people can use motor representations and feedback efficiently to produce pitches accurately when they cannot hear their own voices.

Moreover, musical memory has never been investigated in separate auditory and motor components. Thus addressing the problem of whether memory for music is mainly based on relative or absolute features, the study aims to uncover the nature of musical memory by investigating the extent to which memory for these musical parameters is dependent on an auditory or motor modality.

CHAPTER II

METHODS

Participants

Forty volunteers from the student body of Vanderbilt University and Blair School of Music participated in the study. The participants were between 18 and 27 years of age (mean of 22 years). Of the 40 participants, 12 were **trained-singers (TS)** with a mean vocal training of 4 years, who were recruited from Blair School of Music; 12 were **trained-instrumentalists (TI)** with a mean instrumental training of 6 years and with no history of formal vocal training, who were recruited from Blair School of Music; and the 16 were **non-musicians (NM)** with little or no history of formal vocal and instrumental training, who were recruited from Vanderbilt University.

The instrumentalists had received significant amounts of ear training (mean of 4 years) as a part of their instrumental training, thus they had good aural skills by which they can identify pitches, intervals, chords, rhythms and other basic elements of music; however they did not have the 'specialized vocal training' singers had. This group of people was also particular interest since they had motor representation for a particular pitch based on their motor access to it through the instruments they are playing; however this was in a different way than the trained-singers had.

Apparatus

Participant Recording Sessions

The participant recording sessions took place in a quiet room. Participants wore a set of passive sound isolating earphones (Creative MZ0365 EP-830) through which they heard the masking stimulus; and to further exclude self-produced and environmental sounds they wore a set of active noise-canceling headphones (Bose Quiet Comfort 2) over the earphones. Sound isolating earphones, which were securely fit in the ear canal like an earplug, efficiently blocked-out self-produced auditory feedback while the masking stimulus came straight in.

A microphone (SigmaTel High Definition Audio CODEC) was maintained at a fixed distance of approximately 20 cm from the participants' mouth. Participants' productions were digitally recorded to a laptop computer (Dell Inspiron 1440) using a digital audio editor program (Adobe Audition 3.0) for later analysis.

Masking Stimuli

The masking stimulus consisted of two different kinds of auditory masks: Multitalker (20 talkers) Babble-Mask (Auditec, St Louis, MO, USA) and Song-Mask (the actual song to be sung played at a high volume). Our preliminary studies have shown Babble Mask to be more effective compared to a White-Noise Mask (Band-pass filtered white noise at 50 Hz and 2000 Hz); hence Babble-Mask was used in the study since it has been shown to disrupt performance more reliably than a white noise mask.

Since Happy Birthday is a song that can be played and sung in many different keys, the song-mask was digitally shifted in pitch to match all possible keys, and the participants have been presented with the version that fits their memory representation individually. Song-Mask mimics the situation in which singers cannot hear their own voices due to a loud choir accompaniment; and to our knowledge it has never been used as an auditory mask before. Moreover, the perfect match between the masking frequencies and intended frequencies made it difficult to distinguish between self-produced feedback and feedback provided by the mask. It has been known that the more similar the masking frequencies are to the intended frequencies, the more effective such a mask is at blocking-out self-produced feedback (Blake and Sekuler, 2006).

Procedure

Prior to beginning the experiment, all participants were given a chance for vocal warm up. After the warm up, they were asked to Happy Birthday one time in the key that they preferred. Participants were specifically instructed to imagine that the song was actually playing in their heads before they start singing. If they felt uncomfortable with the stating key, then they were allowed to begin again in a different key (starting with a lower or higher note), but asked to keep the starting note in mind throughout the whole experiment and to try to start from the same tune in each trial. The purpose of this one time singing was to set the key of Happy Birthday (thus the version of Song-Mask) for each participant individually.

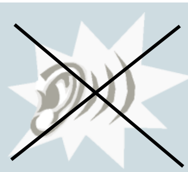
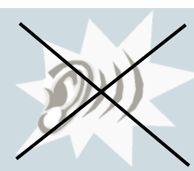
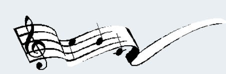
After the dominant key was determined individually, the volume levels for the masking stimulus were determined. In order to identify the optimum volume level for the masking stimulus, the volume was gradually increased while the participants sang Happy Birthday once at each masking conditions, until the point that they could not hear their own voices. The upper volume limit for the masking stimulus was set at 90 DB; but since loudness (subjective impression of intensity) remains a subjective experience that cannot measured by any instrument (Blake and Sekuler, 2006) and since the size of ear canal is different for each person (affecting the actual intensity of the masking stimulus); the threshold for the stimuli was determined on an individual basis. The participants verbally indicated that they could not hear their own voices anymore and the loudness level to achieve this was measured two times for each participant for verification.

Participants were asked to sing at a constant loudness. However, since they could not hear themselves singing and tend to go louder with the masks, in order to assist them in controlling the intensity of their vocal output they were provided with visual feedback. Participants monitored their volume level through a Sound Level Meter (RadioShack 7-Range Analog Display, Model: 33-4050) and the device helped them to keep the vocal productions below 70 DB.

In the experimental conditions, participants sang Happy Birthday repeatedly under two different normal feedback conditions (A Cappella and Sing-Along) and two different masking conditions (Babble-Mask and Song-Mask). Happy Birthday was chosen because of its universal popularity and traditionality; also it was assumed that regardless of musical history everyone had similar amount of past experience singing this song.

The four conditions resulted from crossing two variables: singing from memory vs. singing along with the song; and singing with normal feedback vs. singing with an auditory mask (Table 1). Subjects could hear themselves naturally when singing A Cappella and when singing-along with the recording of Happy Birthday played at a soft volume. Subjects could not hear themselves when singing with the Babble-Mask and when singing along with the recording of Happy Birthday at high volume (Song-Mask). In the sing-along and song-mask conditions, there was reinforcement of the tunes; i.e. subjects were able to hear the actual tunes they were supposed to sing. Each subject sang happy birthday a total of 12 times, 3 times repeatedly in each of four conditions. The order of singing was counterbalanced across subjects so that half of the subjects started with feedback present conditions (a cappella or sing-along) first, and half of the subjects started with memory conditions (a cappella and babble-mask) first, in such a way that 'feedback present' and 'feedback absent' conditions were always presented in alternation.

Table 1. Experimental Conditions

A CAPPELLA	BABBLE-MASK	SING-ALONG	SONG-MASK
			
		 Reinforcement of Tunes	 Reinforcement of Tunes

Analysis

Accessing Pitch Accuracy

In order to identify individual pitch values participants have produced, the frequency analysis function of the digital audio editor program (Adobe Audition 3.0) along with a chromatic tuner (Korg CA-40 Large Display) were used. Two methods have been used simultaneously to assure reliability; and the F0 has been identified in terms of the nearest semitone, i.e. quantized to the nearest pitch (for instance, C4 ± 50 cents = C4). After the notes people sang have been identified, these productions were later compared to the equivalent target notes of the actual song (Figure 2).

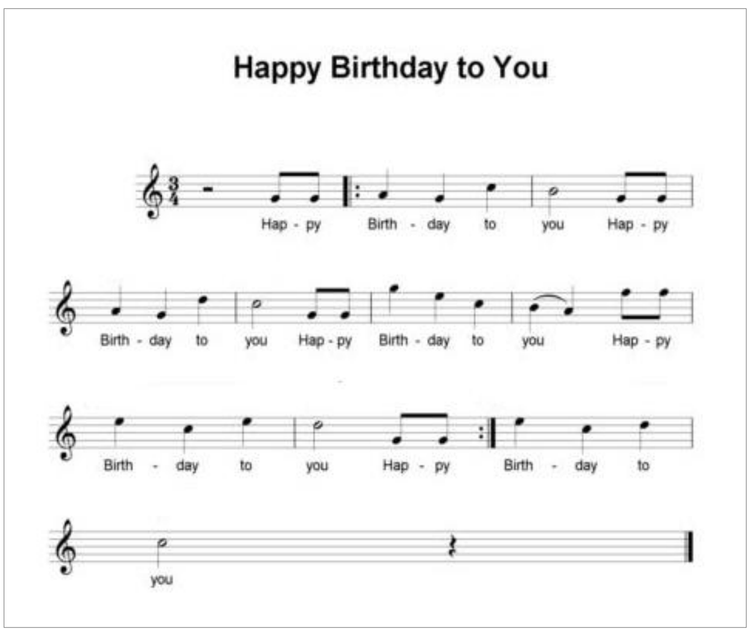


Figure 2. The Score of Happy Birthday in C Major

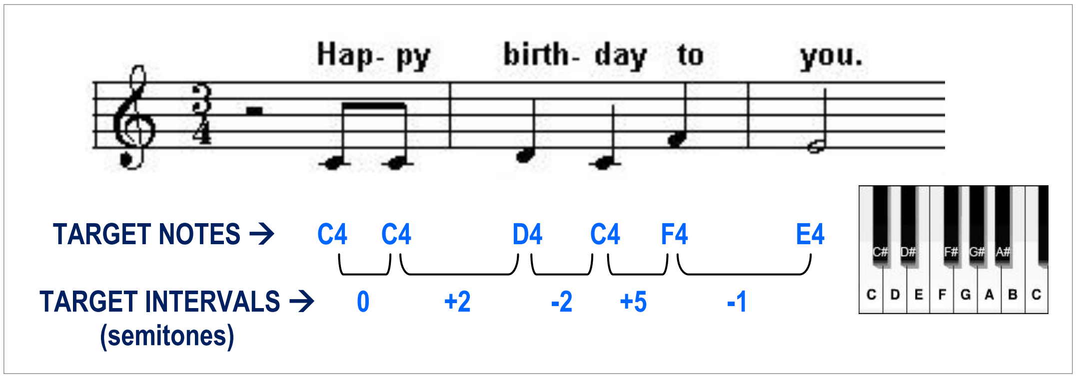


Figure 3. Illustration of Target Notes and Target Intervals

Figure 3 gives the illustration of target notes and target intervals for Happy Birthday sung in C Major. *Pitch errors* were calculated in two different ways: (1) in terms of 'weighted' semitone deviations from the correct "*intervals*" - pitch distances between successive tones (**Relative Measure)** and (2) in terms of 'weighted' semitone deviations from the correct "*individual tones*" (**Absolute Measure)**. First, constant errors were calculated based on semitone differences across produced and target pitches/intervals; and these errors were independent of the direction of movement (singing higher or lower). Then, weighted errors were computed by dividing the constant errors (in semitones) by target intervals (in semitones). These are 'normalized' in the sense that, for instance, a constant error such as 2 semitone error would not be the same for hitting a note which is minor third (m3) above or major sixth (M6) above; because singing a larger interval requires a bigger demand in terms of performance. Thus, weighted method gives a measure of the accuracy of singing by taking into consideration of the relative size of the target interval.

The idea that interval sizes are constant regardless of the key of the song forms the basis of Relative Measure. *Relative pitch accuracy* does not contain information about the actual pitches, and with this measure the pitch errors were calculated on an interval by interval basis. With this respect the relative measure has been adopted as a way to approach *relative pitch memory*. Since the identity and recognizability of a melody is maintained through transposition in pitch, this is the most natural way to access pitch accuracy. *Absolute pitch accuracy*, on the other hand, contains information about the key of the song as well as the actual individual pitches. Thus, the Absolute Measure - a more stringent measure - is sensitive to the key people are supposed to sing in; so with this measure pitch errors were calculated on a pitch by pitch basis. With this respect it has been adopted as a way to approach *absolute pitch memory*. Thus, Relative and Absolute Accuracy were computed in order to access relative and absolute pitch memory representations, respectively.

Accessing Tempo and Rhythm Accuracy

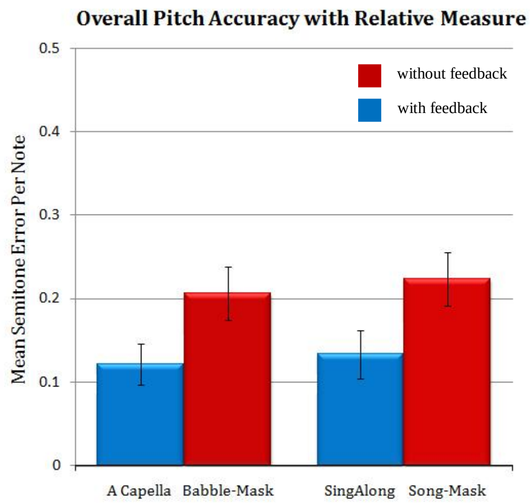
The duration of each subject's production per note (in sec) has been identified using a time keeper and these measurements were accurate to within 1 msec. *Tempo accuracy* was accessed by comparing the tempo of each subject's productions across A Cappella and Babble-Mask conditions, for which the participants had complete control over the timing of their singing. For calculating tempo, the total selection was divided into units of beats yielding duration per meter; and the data were presented as tempos in units of ‘beats per minute’. *Rhythmic accuracy,* on the other hand, was accessed by comparing the “relational” durations of each note across the two conditions. Unlike tempo accuracy, rhythmic accuracy was measured based on the relationship (ratio) between the lengths of two successive notes and accessed by comparing the durational ratio of two successive notes across A Cappella and Babble-Mask conditions.

CHAPTER III

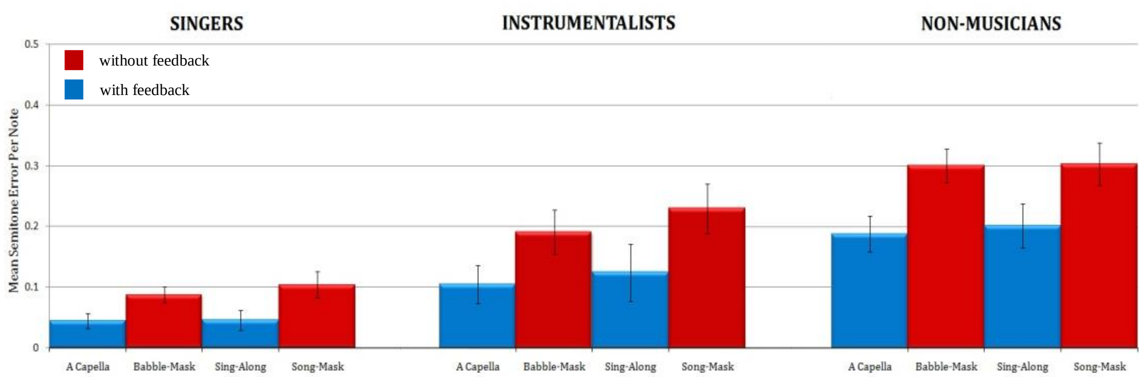
RESULTS

Pitch Accuracy

Statistical analysis was completed on data calculated via both relative and absolute measure, using a three-way (2x2x3) repeated measures analysis of variance (ANOVA) with two within-subject factors and one between-subject factor. The within-subject factors were masking (with/without auditory feedback) and memory (memory/model based), the between-subject factor was group (TS, TI and NM). With the relative measure, which has been taken as the most natural way to access pitch accuracy, ANOVA revealed a significant main effect for feedback (F = 157.362, p < .00) a main effect for group (F = 53.044, p < .00) and a significant interaction effect between feedback and group (F = 6.493, p < .004), with no significant main effect for the model ('singing from memory' vs. 'singing along with the song') (F = 3.511, p > .069). Post-hoc testing utilizing an LSD Fisher test, which was computed on group means for accessing overall pitch accuracy, revealed that all groups differed from one another significantly at p<.00 (TS - TI), p<.00 (TS – NM) and p<.00 (TI - NM). When taking into account of overall pitch accuracy regardless of the condition, TI were most accurate, NM were least accurate, whereas TI were in between. Figure 4 illustrates pitch accuracy for four different conditions with relative measure (a) for the average overall population and (b) on a group basis. These plots are also of interest as they show that the TS group was not only most accurate with and without feedback, but also was less variable within the group. TI and NM groups were more variable with greater spread around the mean compared to TS.



a)



b)

Figure 4. Pitch Accuracy with the Relative Measure across Four Conditions (a) for the Average Overall Population (b) on a Group Basis. The blue bars represent conditions in which people have sung 'with feedback' and red bars indicate conditions 'without feedback'. Y axis indicates the mean semitone errors per note/interval.

A difference score was obtained for each response by subtracting errors in the 'feedback absent' conditions with errors in the 'feedback present' conditions. Because of the main effect for feedback and the interaction effect between feedback and group, group means were compared on 'feedback present' and 'feedback absent' conditions separately, and of particular interest, on difference scores using three separate one-way ANOVAs. Results of the one-way ANOVAs revealed a significant effect for group for 'feedback present' conditions (F = 27.186, p < .00), for 'feedback absent' conditions (F = 68.728, p < .00) as well as for 'difference scores' (F = 6.494, p< .004). Post-hoc paired comparisons (LSD) revealed that each group mean differed from one another significantly at (p<.003 (TS - TI), p<.00 (TS – NM), p<.00 (TI - NM) in 'feedback present' conditions, and at p<.00 (TS - TI), p<.00 (TS – NM), p<.00 (TI - NM) in 'feedback absent' conditions. TS were most accurate when both singing with and without auditory feedback compared to TI and NM, NM were least accurate for both cases, whereas TI were in between. As a follow up on the interaction effect between group and feedback observed in the main ANOVA, pairwise comparisons on differences scores were computed. Results revealed that TS differed from NM (p <.001) and TI (p < .013) significantly, whereas no significant difference was observed between TI and NM groups (p > .491). That is, TS group was less affected by the absence of auditory feedback compared to both TI and NM, whereas no such difference was observed for TI and NM groups, i.e. absence of auditory feedback for TI was as disruptive as it was for NM. Figure 5 shows this effect in a more clear way.

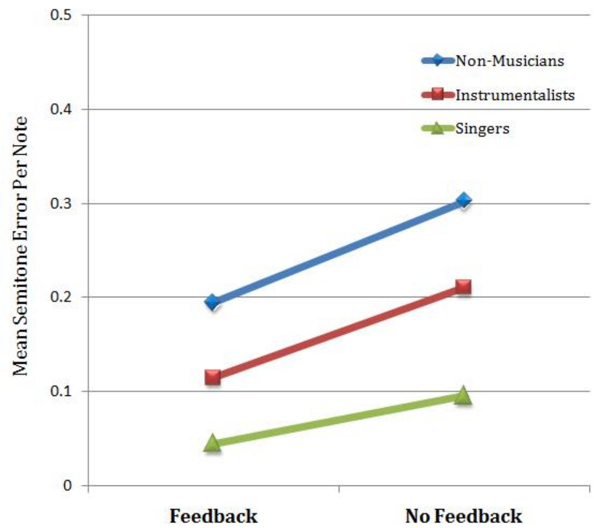


Figure 5. Pitch Accuracy with the Relative Measure across ‘feedback’ and ‘no feedback’ Conditions for 3 Different Groups. The y axis represents mean semitone error per note/interval.

These results can also be interpreted in proportional terms, for which TS’ pitch errors in ‘feedback absent’ conditions appear twice as big as they are for ‘feedback present’ conditions; whereas for NM this effect is only 3/2 of the normal feedback condition. Such result could have been arisen due to the ‘floor effect’ observed in TS group. TS were already very good at pitch accuracy without auditory feedback, so they had less chance to improve performance when singing with feedback. NM, on the other hand, had more errors to start with, so they had more chance to improve performance when singing with feedback. Interpreting data in term of ratios would also suggest that TS were better able use auditory feedback compared to NM since they had more benefit of auditory feedback; nevertheless magnitude of improvement was still bigger for NM and perceptually singers always sounded better than NM when sang with and without feedback. Hence, the magnitude of difference scores was more instructive for explaining the size of the effect for absence of feedback.

Thus using magnitude of error –rather than ratios– was more meaningful and suitable when accessing pitch accuracy.

Of particular interest was that with the relative measure there was no significant effect of presence of memory/model for either group; that is, it did not made a difference whether people sang from memory or sang along with the recording. This result was surprising, because based on post experimental questionnaire, %85 of the participants told that they found Babble-Mask to be more difficult than the Song-Mask for accurate singing. However, the results have shown Song-Mask to be as effective as Babble-Mask in disrupting performance, even though people received constant additional auditory reinforcement.

Three-way (2x2x3) repeated measures analysis of variance (ANOVA) completed on data gathered through Absolute Measure revealed a main effect of feedback only (F= 38, p <.009). Figure 6 shows pitch accuracy for overall population with two different measures (relative versus absolute measure). Post-hoc paired *t*-tests on each pair of each measure showed that with the exception of memory based pair of absolute measure, all pairwise comparisons between ‘feedback present’ and ‘feedback absent’ conditions were significant at p<.00. The big variability observed in memory based pair of the absolute measure resulted from the fact that absolute measure was sensitive to the key in which the song was sung; and participants were not able to stay within the same key; they tended to shift between keys throughout the repetitive trials. Thus, those conditions did not provide a reliable way to compare the two measures. Model based conditions, on the other hand, were experimentally ‘controlled’ since there was constant imposition of key and participants were assured to stay within the original key. These two conditions promised a better estimate for the difference between absolute and relative pitch accuracy.

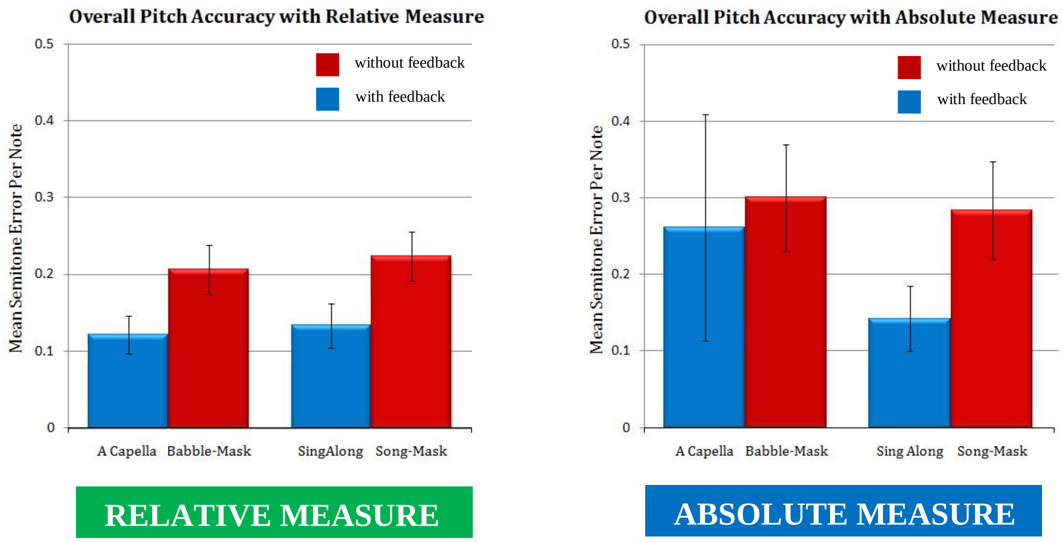
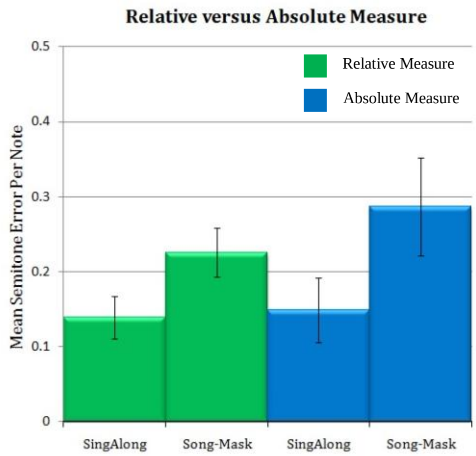


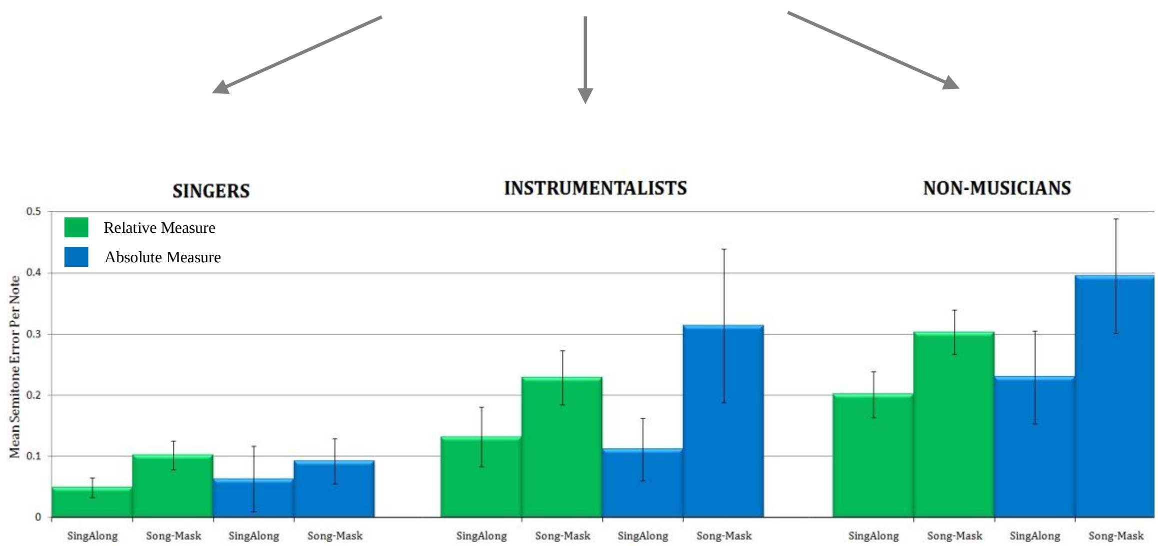
Figure 6. Average Overall Pitch Accuracy across Four Conditions with Relative versus Absolute Measure. The graph on the left shows errors with the Relative Measure, and the graph on the right shows errors with the Absolute Measure. Y axis indicates mean semitone errors per note/interval.

In order to access relative and absolute pitch accuracy, another statistical analysis was carried out on overall data on the ‘singing along with the song’ conditions by means of a three-way (2x2x3) repeated measures ANOVA, with masking (with/without feedback) and ‘measure type’ (relative/absolute) as the within-subject factors and ‘group’ (TS, TI, NM) as the between-subject factor. Two subject’s data have been excluded from the calculations since the subjects have reported to possess *absolute pitch*. There was a main effect of measure ($F = 4.861$, $p < .034$), a main effect of feedback ($F = 71.929$, $p < .00$), a significant interaction effect for feedback and group ($F = 6.521$, $p < .004$) and of particular interest there was a significant interaction effect of ‘measure type’ and ‘feedback’ ($F = 6.486$, $p < .015$). The interaction indicated that the difference between ‘with feedback’ and ‘without feedback’ conditions were more pronounced with the absolute measure compared to relative measure; a result which implied that absolute pitch memory was more dependent on auditory feedback.

Pitch accuracy with relative versus absolute measure on model based conditions (a) for the average overall population and (b) on a group basis have been shown in Figure 7.



a)



b)

Figure 7. Pitch Accuracy with Relative versus Absolute measure (a) for the Average Overall Population (b) on a Group Basis. The green bars represent pitch errors with Relative Measure, and the blue bars indicate pitch errors with Absolute Measure. Y axis indicates the mean semitone errors per note/interval.

Because of the interaction, three separate two-way ANOVAs with masking (with/without feedback) and measure type (relative/absolute) as the within-subject factors were computed for each group on the model based conditions. The results indicated a significant interaction effect for measure type and feedback for TI ($F = 5.975$, $p < .035$) and NM ($F = 5.741$, $p < .029$); whereas no such interaction was found for TS ($F = .607$, $p > .454$). The absence of auditory feedback did not affect absolute accuracy for TS as they did for TI and NM, indicating that absolute pitch accuracy (compared to relative) was more dependent on auditory feedback **only for** TI and NM.

The data have also been analyzed to see whether or not there was a trial/training effect resulted from repetitive singing. A two-way ANOVA with 'trial number' as the within-subject factor and 'group' as the between-subject factor has been computed for each measure. Throughout the 3 repeated trials, no significant change of intonation accuracy was observed for either group with either measure, suggesting that a total of 12 times singing was not enough to improve the use of either feedback mechanism (Figure 8).

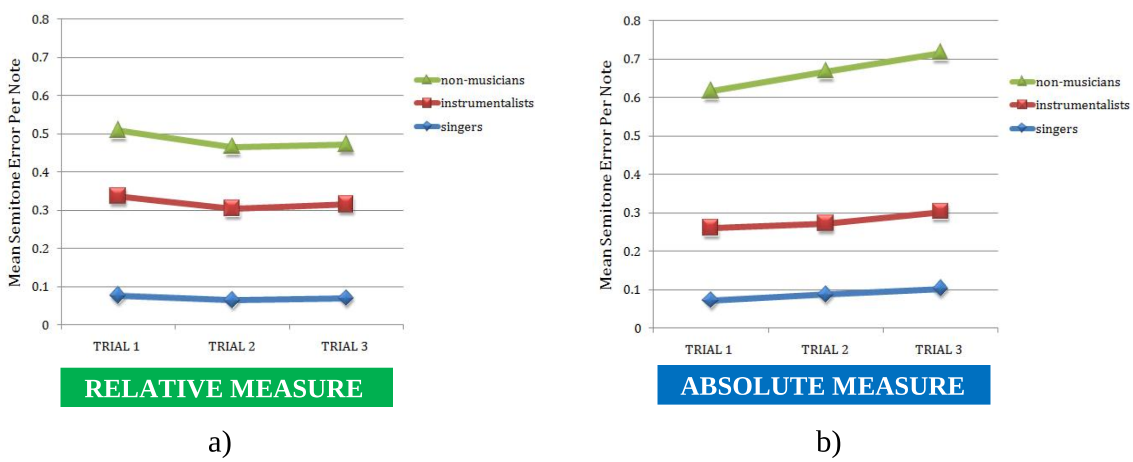


Figure 8. Trial Effect on a Group Basis with (a) Relative Measure and (b) Absolute Measure. The X axis represents trial numbers independent of conditions.

Tempo and Rhythm Accuracy

For the overall data, a paired sample t-test was computed on tempo scores for A Cappella and Babble conditions, for which the participants had complete control over the timing of their singing. As shown in Figure 9 a significant difference was found between the unmasked and masked conditions (t = 4.683, p < 0.00). When auditory feedback was blocked-out, participants on average slowed down significantly. Rhythmic ratios, on the hand, which were computed on an interval (two successive notes) basis through a paired sample t-test, did not show any significant difference between the masking conditions. Figure 10 shows how the same rhythmic pattern was preserved across two different conditions despite changes in tempo. Unlike tempo, rhythmic pattern was preserved across different conditions of masking, a result which indicated that rhythmic accuracy is successfully maintained through the motor system without need for auditory calibration.

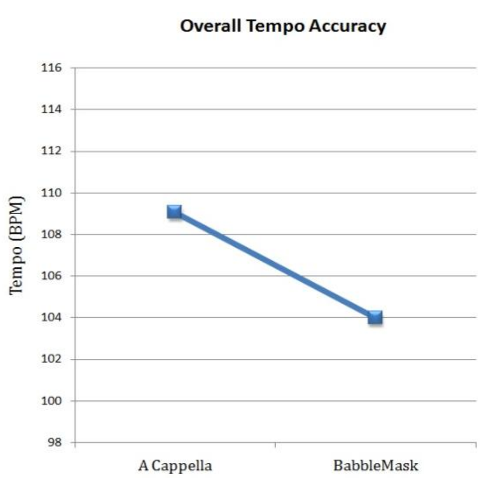


Figure 9. Average Overall Tempo Accuracy across A Cappella and Babble Mask Conditions The y axis represents Tempo in beats per minute.

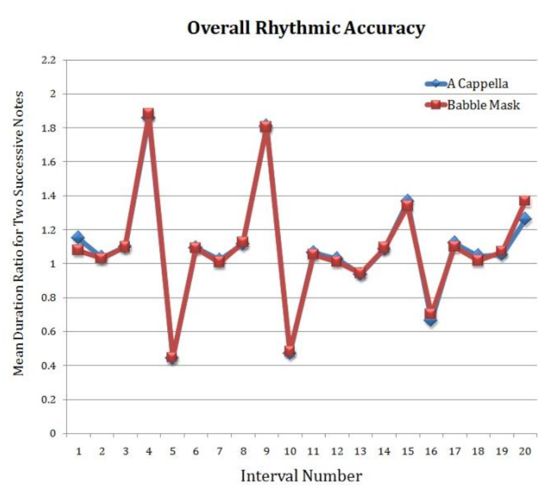


Figure 10. Average Overall Rhythmic Accuracy across A Cappella and Babble Mask Conditions The y axis represents Tempo in beats per minute.

In order to access group based differences, a two-way repeated measures ANOVA with masking (with/without feedback) as the within-subject factor and group (TS, TI, NM) as the between subject factor was performed. The data for 7 out of 40 subjects, which have been reported to fasten in the babble-mask condition, in contrast to the general trend, have been excluded from this calculation. The results revealed a main effect of feedback (F = 54.290, p <.00) and a marginally significant group x masking interaction (F = 2.747, p<.08). Post-hoc testing utilizing an LSD fisher test on difference scores between masking conditions revealed that the tempo of the TS group was less effected by the absence of feedback compared to TI and NM groups (p < .029) whereas no difference was observed between TI and NM groups in terms of the effect of masking. Figure 11 shows how TS on average slowed down less compared to other two groups. Rhythmic stability across two conditions, on the other hand, has been further confirmed on a group basis. Rhythmic ratios, unlike tempo, did not show any significant difference between the unmasked and masked conditions on a group basis either. Figure 12 shows how each group preserved rhythmic accuracy despite group based differences on tempo accuracy. So, similar to the idea that absolute pitch accuracy for TS (unlike other groups) was more stable with absence of auditory feedback; again their absolute accuracy for time (tempo) was more stable with absence of auditory feedback compared to other groups.

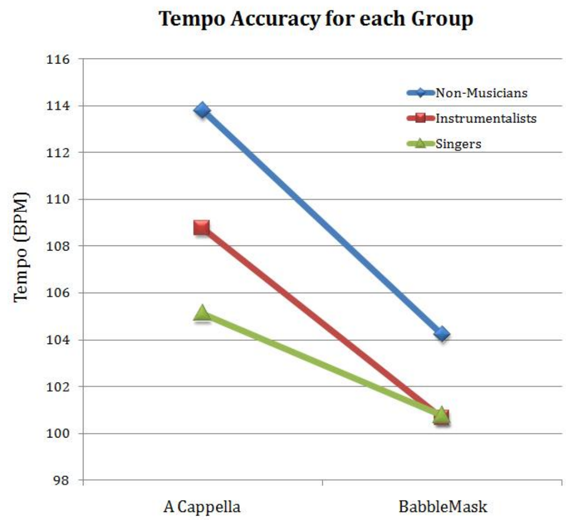


Figure 11. Tempo Accuaracy on a Group Basis across A Cappella and Babble-Mask conditions. The y axis represents Tempo in beats per minute.

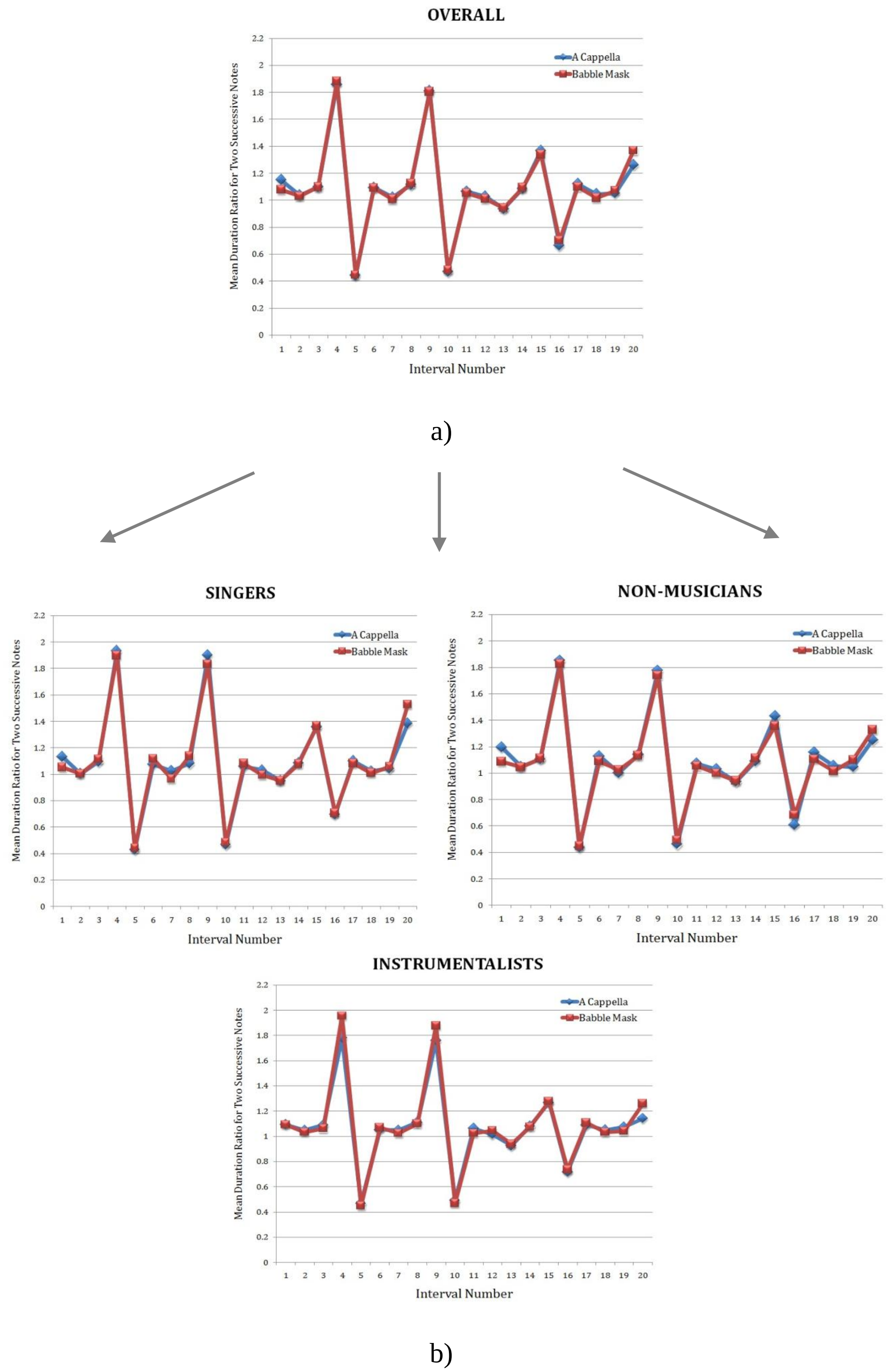


Figure 12. Rhythmic Accuaracy (a) for the Average Overall Population and (b) on a Group Basis. The X axis represents Interval Number, and the Y axis represents mean durational ratio for two successive notes.

CHAPTER IV

DISCUSSION

Results showed that intonation accuracy was reduced without auditory feedback for all three groups. That is, when people were not able to hear their voices (through a babble or song-mask) their intonation accuracy was worse. This suggests that 'auditory feedback' acts as an important factor for maintaining intonation accuracy in general, and it remains important even for students of voice who have been taking vocal training for approximately 4 years. Given that musical expertise requires years of dedication and practice, it is likely that auditory feedback remains significant even after 4 years of training. Our results have also shown that TS were most accurate in pitch with and without feedback, NM were least accurate, whereas TI were in between. Thus, ear training and instrumental education TI had helped them to maintain overall pitch accuracy. However, our results have also shown that TS's performance was less disrupted by the absence of auditory feedback compared to other groups, whereas for TI the absence of auditory feedback was as disruptive as it was for NM. That is, TI was not different from the NM in terms of the magnitude of disruption when singing without auditory feedback. Thus, although TI was better than NM in terms of overall pitch accuracy, ear/instrumental training did not help them to use motor representation and feedback effectively in cases where auditory feedback could not be utilized. Singers, on the other hand, were more resistant to the effects of masking (less disrupted in masking conditions) and were able to use motor/proprioceptive information more efficiently when auditory feedback was not available. Because of the kind of training singers are receiving, which focuses on the sensations of notes rather than hearing them, they have improved ability to use motor information to control pitch, and they do not need to rely on auditory information as much as other people do. Indeed, in our sample of trained singers there was one participant with 10 years of professional individualized vocal training who is active in formal singing (e.g., school choirs and solo recitals). She has shown almost no disruption of performance when auditory feedback was blocked-out. Similarly Murbe et al. (2004) in a longitudinal study has shown that after three years of professional schooling, the singers showed improved performance under conditions of no auditory feedback. Thus, accurate use of motor feedback for the control of pitch is subject to the level of training of the singer, with more training yielding better ability to control pitch without auditory feedback.

On the other hand, there were also two subjects in the NM group who have not shown any kind of disruption in pitch accuracy when auditory feedback was not available. For one of these subjects her overall performance was aligned with the NM group performance on overage, whereas other subject's performance was aligned with TS group performance. Some individuals, then, despite any background in vocal training and/or formal singing (church, school choirs) seem to naturally use motor feedback more efficiently in cases when auditory feedback is absent. Similarly, Watts et al. (2003) has found that although singers are more resistant to the effects of masking than non-singers, talented non-singers perform even better than trained singers when auditory feedback is unavailable. Thus, it should be noted that the effects of training and natural talent may be confounded.

The results have also shown Song-Mask to be as effective as the Babble-Mask in blocking-out auditory feedback. This result was surprising since %85 of the subjects found song-mask to be easier in comparison to babble-mask for accurate singing. However, the perfect frequency match between the masking frequencies and self-produced frequencies affirms song-mask as an effective masking stimulus. It has been known that noise is very affective as a mask when it contains frequencies ‘at’ or ‘near’ the frequencies of the test tones (Blake and Sekuler, 2006). Hence, in this case, the perfect frequency match between the masking frequencies and self-produced frequencies made it difficult to distinguish between self-produced feedback and feedback provided by the mask. In other words, the expected auditory feedback was imitated by an external producer and externally presented auditory information was perfectly in line with the expectations of the motor system. Thus, this might have fooled the vocal motor system, leaving subjects with less focus on the available feedback mechanism; namely the proprioceptive feedback; and giving them the illusion of they were in tune when singing along with the Song-Mask.

A related finding was that the presence of the song (the reinforcement of the actual tunes to be sung) did not make a difference for intonation accuracy. That is, the performance scores on memory based and model based conditions were similar regardless of the group. Although, there is no clear implication for this finding, it suggests that either people cannot efficiently use the external auditory information, or Happy Birthday is such a practiced tune that it does not require extra attention to the original tunes.

The conditions in which people sang along with the song provided a controlled situation since subjects were assured to stay within the original key. For these conditions, for the general population, the effect of the mask was more pronounced with absolute measure compared to the relative measure, suggesting that absolute pitch accuracy was more disrupted by the absence of auditory feedback. Relative pitch accuracy, on the other hand, was less disrupted by the absence of auditory feedback, implying that relative pitch memory is more readily handled by the motor system, i.e. more dependent on a motor component. In other words, in the cases where auditory feedback is limited, people on general can access relative pitch information more easily via the motor system, whereas absolute accuracy requires more incorporation from the auditory modality.

On a group basis similar results held for TI and NM; whereas for TS the reliance on auditory feedback for absolute accuracy was lost. When singers were not able to hear their own voices, their absolute accuracy was not as affected as it was for TI and NM. In other words, the dependence on auditory feedback for absolute accuracy has been reduced in singers; and they could access the absolute information more readily than other groups under conditions of no auditory feedback. This suggests that with vocal training, which focuses on placement of notes in an absolute manner, singers gain increased incorporation of the absolute information into the motor system. These findings are in line with professional vocal training which emphasizes placement of individual notes. For instance, there is something specific about what a particular note, such as F5 in soprano voice feels internally during vocal performance. The voice students need to memorize that particular sensation so that when the teacher gives them a piece that either starts on F, ends on F, F in the middle of the line, regardless of the vowel, they have to be able to recognize that that is the place they can feel the note. This is, in a way, similar to gaining *Absolute Pitch* (a rare phenomenon found in 1 of 10000 people) in the **motor** domain. Thus, through training, absolute pitch memory becomes less dependent on an auditory component; absolute pitch memory becomes more "motoric".

The findings on tempo and rhythmic accuracy were parallel to what was found for pitch accuracy. In the absence of auditory feedback, there was a significant decrease in tempo; that is, on average people slowed down when they did not hear themselves. This implies that the motor plan has already incorporated predictions with respect to auditory feedback and the motor loop (thus people's actions) slow down due to this violation of expectation. However, even if people slow down when auditory feedback is blocked-out, they do still preserve rhythmic accuracy. This suggests that tempo-free rhythmic information lies deep into the motor system, and rhythmic calculations are more based on an internal motor system. Addressing the idea that tempo is described most naturally in absolute terms, whereas rhythm is more about the relationship between notes, hence relative; the absolute time parameter – tempo – (compared to the relative time parameter) is more affected by the absence of auditory feedback. Thus, for the general population, the relative information (also) associated with musical time is more readily handled by the motor system. In other words, similar to what has been found for the absolute pitch accuracy for overall population; absolute time parameter tempo also needs more auditory calibration; whereas, relative time parameter 'rhythm' is more readily handled by the motor system, as is relative pitch memory.

On a group basis, on the other hand, singers have been found to slow down less compared to TI and NM. For singers the reliance on auditory feedback for absolute accuracy in musical tempo has been reduced due to extensive training they were receiving. Nevertheless, each group preserved tempo-free rhythmic accuracy when they lacked auditory feedback, providing further confirmation that rhythmic accuracy is successfully maintained through the motor system without need for auditory calibration. Thus, similar to the idea that singers' absolute memory for pitch was less affected by the absence of auditory feedback compared to other groups; again their absolute memory for time (tempo) was less affected by the absence of auditory feedback, suggesting a more stable absolute memory representation compared to other groups. Thus, in singers the absolute pitch and time (tempo) is more readily handled by the motor system. For singers, the absolute accuracy in pitch and time relies less on auditory feedback and is more deeply encoded in the motor system, i.e. they can access this information more readily when auditory feedback is not available.

In sum, for the general population, absence of auditory feedback disrupts intonation accuracy significantly, and when auditory feedback mechanism is blocked out, people can access relative information (in comparison to the absolute) of pitch and time more readily, due to its higher dependence on a motor modality. This also suggests that memory for this particular familiar song is more stable for melodic intervals and tempo-free rhythmic information than is memory for the absolute pitch and tempo, due to their higher dependence on the motor system.

For trained singers, on the other hand, in the absence of auditory feedback their intonation accuracy is less disrupted than instrumentalists and non-musicians. This suggests that vocal training improves the ability to use motor information. Moreover, for singers the reliance on auditory feedback for absolute accuracy is reduced. For trained-singers, compared to other groups, absolute pitch and time (tempo) accuracy is less disrupted by the absence of auditory feedback. That is, when auditory feedback is blocked-out, trained singers can access memory for absolute pitch and time more readily than non-musicians and instrumentalists via the motor system due to the extensive training they are receiving which focuses on placement of individual notes.

Implications and Future Directions

For people with hearing loss and/or speech impediments, a special speech therapy program which focuses on internal sensations might offer an effective solution. The study also provides promising directions for the challenges singing faces by people with severe hearing loss and with cochlear implants.

Continuing research also aims to explore the performance accuracy of people with Absolute Pitch (AP) when singing popular songs that are typically encountered in only one version by a particular artist/group, so that the song is always heard and played in the same key and at the same tempo. People with AP, also called pitch experts, have a remarkable ability to label and produce tones without reference to any other tones, and they are thought to have a solid internal representational template for pitches. It is a rare ability that reportedly occurs in only about 1 in 10,000 people (Takeuchi & Hulse, 1993). Hence, of particular interest, another group of AP possessors are to be tested under masked feedback conditions. One study suggests that true AP possessors do rely somewhat on multiple representations including kinesthetic imagery, for pitch identification (Zatorre et al., 1989); however, no research to this date has tested people with AP under restricted feedback conditions and examined the possibility that representations for pitch in long-term memory may actually be strongly dependent on a motor component. The two different measures (absolute vs. relative) designed to access absolute vs. relative pitch accuracy and discussed in this paper, might yield the difference for AP population.

REFERENCES

Appelman, D. R. (1986). The science of vocal pedagogy: theory and application. Bloomington, Indiana: Indiana University Press. pp. 434.

Blake, R. and Sekular, R. (2006). Perception (6ed.). New York, NY: McGraw-Hill.

Burnett, T. A., Freedland, M. B., Larson, C. R., Hain, T. C. (1998). Voice F0 responses to manipulations in pitch feedback. *J Acoust Soc Am* 103:3153–3161.

Cowie, R., Douglas-Cowie, E. (1992). Postlingually acquired deafness. Trends in linguistics, studies and monographs. Mouton de Gruyter, New York.

Deutsch, D. (1978). Octave generalization and melody identification. *Perception & Psychophysics*, 23, 91-92.

Deutsch, D. (1999). The processing of pitch combinations. In D. Deutsch (Ed.), *The psychology of music* (2nd ed., pp.349-411). San Diego: Academic Press

Di Carlo, N. (1994). Internal voice sensitivities in opera singers. *Folia Phoniatrica et Logopaedica* 46:79–85.

Elliot, L., Niemoeller, A. (1970). The role of hearing in controlling voice fundamental frequency. *Int Audiol 9*:47–52.

Jones, J. A. & Keough, D. (2008). Auditory-motor mapping for pitch control in singers and nonsingers. *Exp Brain Res.* Sep; 190(3):279-87.

Jones, J. A,. Munhall, K. G. (2002). The role of auditory feedback during phonation: studies of Mandarin tone production. *J Phon 30*:303–320.

Levitin, D. J. (1994). Absolute memory for musical pitch: Evidence from the production of learned melodies. *Perception & Psychophysics* 56(4), 414-423.

Levitin, D. J. & Cook, P. R. (1996). Memory for musical tempo: Additional evidence that auditory memory is absolute. *Perception & Psychophysics* 58(6), 927-935.

Leder, S. B., Spitzer, J. B., Kirchner, J. C. (1987). Speaking fundamental frequency of postlingually profoundly deaf adult men. *Ann Otol Rhinol Laryngol* 96:322–324.

Monahan, C. B. (1993). Parallels between pitch and time and how they go together. In T. J. Tighe & W. J. Dowling (Eds.), Psychology and music: The understanding of melody and rhythm. Hillsdale, NJ: Erlbaum.

Murbe, D., Pabst, F., Hofmann, G., & Sundberg, J. (2002). Significance of auditory and kinesthetic feedback to singers' pitch control. *Journal of Voice,* 16(1), 44-51.

Murbe, D., Pabst, F., Hofmann, G., & Sundberg, J. (2004). Effects of a professional solo singer education on auditory and kinesthetic feedback – a longitudinal study of singers' pitch control. *Journal of Voice*, 18(2), 236-241.

Oller, D. K., Eilers , R. E. (1988). The role of audition in infant babbling. *Child Dev 59*:441–449.

Plantinga, J., & Trainor, L. J. (2005). Memory for melody: Infants use a relative pitch code. *Cognition,* 98, 1–11.

Raphael, LJ, Borden, GJ, & Harris, KS. (2007). Speech science primer. 5th ed. Baltimore: Lippincott Williams & Wilkins; pp. 166-76.

Roebuck, J. (29 September 2007). I am a deaf opera singer. In Guardian.co.uk. Retrieved June 9, 2010 from http://www.guardian.co.uk/theguardian/2007/sep/29/weekend7.weekend2

Schellenberg, E. G., & Trehub, S.E. (2003). Good pitch memory is widespread. *Psychological Science,* 14, 262–266.

Serafine, M. L. (1979). A measure of meter conservation in music, based on Piaget's theory. *Genetic Psychology Monographs*, 99, 185-229.

Takeuchi, A.H., & Hulse, S.H. (1993). Absolute pitch. *Psychological Bulletin, 113(2),* 345-361.

Trainor, L. J., McDonald, K. L., & Alain, C. (2002). Automatic and controlled processing of melodic contour and interval information measured by electrical brain activity. *Journal of Cognitive Neuroscience,* 14, 430–442.

Ward, W. D., & Burns, E. M. (1978). Singing without auditory feedback. *Journal of Research in Singing & Applied Vocal Pedagogy*, 1, 24-44.

Watts, C., Murphy, J., & Barnes-Burroughs, K. (2003). Pitch-matching accuracy of trained singers, untrained subjects with talented singing voices, and untrained subjects with non-talented singing voices in conditions of varying feedback. *Journal of Voice*, 17, 185–194.

Xu, Y., Larson, C., Bauer, J., Hain, T. (2004). Compensation for pitch-shifted auditory feedback during the production of Mandarin tone sequences. *J Acoust Soc Am* 116:1168–1178.

Zatorre, R. J., & Beckett, C. (1989). Multiple coding strategies in the retention of musical tones by possessors of absolute pitch. *Memory & Cognition, 17(5),* 582-589.