Hearing, Feeling, and Performing: Masking Studies with Trombone Players

Perry R. Cook
Department of Computer Science,
Department of Music, Princeton University

This work is motivated by a desire for understanding the use of feedback and feedforward control mechanisms in musical performance; specifically, to discover the role of each of the possible control mechanisms that the skilled musician has available and uses in performance. In nearly all instruments, it is likely that there are control components of a musical performance which exploit all of the different functions of memory, auditory feedback, and various touch feedbacks via sensors of position, skin contact pressure, and vibration. The nature of each instrument and instrument family, however, dictate that different mechanisms will dominate, and the skilled instrumentalist will adopt control strategies which yield the greatest control and success in expressive musical performance. Further, pedagogy for a given instrument family will also adopt teaching and rehearsal practices which develop and reinforce the most successful control mechanisms.

Studies of the effects of auditory masking on singers have been conducted (Campbell & Michel 1980; Elliot & Niemoeller 1970; Shipp et. al. 1984; Sundberg 1981; Ward & Burns 1978; Ternström et. al. 1983), showing that even singers with perfect pitch can make gross errors in pitch production when auditory feedback is impaired. Further, singers when presented with a delayed version of their own voice demonstrate changes in long-term and short-term pitch control (Deutch & Clarkson 1959; Shipp et. al. 1984). Very little research on auditory masking and performance has been done using instrumental performers. The role of touch and vibration sensation has been investigated in bowed strings and piano (Askenfelt & Jansson 1992). The study described in this paper examines the role of feedback control in trombone playing. The trombone is an instrument which allows the player to exert continuous control over pitch via the smooth motion of the slide, and the player has intimate contact with the instrument via the lips, which serve as the non-linear oscillating "reed" in brass instruments. Traditional brass pedagogical literature (Kleinhammer 1963) notes that the principal mechanism for tonal "centering" is the "feel" of the note, that is, once the slide position is set, small adjustments of the lip tension bring the lip natural frequency into alignment with the nearest impedance peak of the trombone bore resonator. Pedagogical literature from circa 1850 discusses the penalty for "lipping" notes away from the natural frequencies of the bore (Arban 1936), and suggests that player feeling is critical to good technique.

The level of auditory masking used in this study was 110 dB, A-weighted, presented via closed headphones for no longer than 30 seconds at a time. Given the average long-term exposure of symphony orchestra players of 76-102 dBA L_{eq} (averaged A-weighted sound level) found by (Royster 1989), the levels used in this study were judged not only safe but typical for many professional musicians. Royster further found that symphony orchestra players exhibited hearing thresholds better than the average population, except for an asymmetrical excess loss in the left ears of violin players, even though the measured sound levels would predict hearing loss greater than the non-industrial noise-exposed population. Some theoretical and experimental reason for this lack of damage due to prolonged exposure might be related to the stapedius reflex, which is a mechanism which attenuates the vibration amplitude of the middle ear (Møller 1972). Little is known about bone conduction in brass instrument players, but using data from (Tonndorf 1972), and data on the masking effects of singer's own voices from (Lindqvist-Gauffin & Sundberg 1974), it can be assumed that the pink noise masking at 110 dBA was highly likely to have masked all auditory feedback from the trombone to the player. Finally, all subjects reported that they could not hear their performance, and were essentially playing "by feel." This will be discussed in more detail later in the paper.

Five trombone players were recruited as subjects for the experiment. Subjects ranged from 15 to 47 years of musical experience, and 7 to 45 years trombone playing experience. All subjects were paid for the experiment. Subjects were asked to learn and perform three simple exercises, which were composed specifically for this study. The first is a basic brass warm-up exercise, requiring normal amounts of lip and slide motion. The second exercise requires less adjustment in lip tension, but large excursions in slide position. The third exercise requires very little slide motion, but larger adjustments in lip tension. Figure 1 shows the three exercises, with slide positions noted below each note. Trombone slide positions range from 1 which is the shortest bore length with the slide positioned against the mouthpiece end, to position 7 with the arm and slide fully extended.

Figure 1 Exercises for trombone auditory masking studies. Slide positions are shown below each note. The first exercise is a basic warm-up, the second uses mostly slide motion, and the third uses mostly adjustments in lip tension. The arrows and brackets mark notes used for specific analyses described later.

Subjects were allowed to rehearse each piece until they were confident of their ability to play it. Given the experience of the players and the simple nature of the exercises, most subjects rehearsed each example at most three times before announcing they were satisfied. Subjects were then asked to play the piece with a metronome at 72 bpm, then at 96 bpm. The subjects were then asked to play the exercise again at 72 bpm, then again at

72 bpm with pink noise auditory masking wearing closed headphones. The non-masked/masked pair was then repeated at 96bpm. All subjects performed these same tasks in the same order, except in the final exercise the three more expert players were asked to sight read twice with auditory masking in place. Those subjects were then allowed to rehearse and perform the third example as with all other examples. The trombone players with transposition trigger valves were asked to play all exercises without using the triggers. All sessions were audio taped and video taped. It is anticipated that the video tape will be used in later studies involving targeting studies for the trombone player arm/slide system.

For quantitative analysis of the performances, first a comparison of missed notes was performed across a total of 136 performances (3099 notes). The average number of missed notes in non-masked performances was 2.07%, with a standard deviation of 1.22%. The average number of missed notes in masked performances was 3.0%, with a standard deviation of 2.20%. Thus, there were 0.93% more missed notes average in masked performances, but this can hardly be considered statistically significant. All subjects played louder in masked performances. The average increase in average intensity with masking was 3.75 dB (StdDev=2.611dB).

Frequency was calculated by transferring the digitally recorded performances directly to computer, down sampling to 8KHz, applying a Hamming window of length 512, Fourier Transforming with zero padding to length 1024, and performing parabolic interpolation of the fundamental peak. Frequency calculations were done at 30 ms. intervals throughout the entire analyzed performances. The parameters of the frequency detection yield estimates with a worst-case error of 0.4% on the lowest notes performed (Brown 1996). All frequency extraction data was inspected for consistent differences in masked/non-masked performances. Notes were aligned by their begin times, and differences were formed between the masked and non-masked fundamental frequencies on a note-by-note basis. Longer steady-state cadence tones were selected from each non-masked/masked exercise pair for comparison of fundamental frequency. These notes are marked with arrows in Figure 1, and the frequency trajectories are shown in Figure 2.

Average frequency and standard deviations of frequency were compared in the long tones of the non-masked/masked performances. Most notes yielded smaller standard deviations on masked performances than on non-masked. This could be explained by some traces of Figure 2 (Subject 2, Slide G, for example), where clear vibrato is present on unmasked performances, but is less visible on masked performances. Otherwise, no consistent differences were found. To study dynamic behavior, segments of the lip and slide exercises which required the most dexterity were analyzed. These sections correspond to the bracketed notes in Figure 1. The frequency trajectories are shown in Figures 3 and 4. Most notable is the profound similarity between the frequency traces of the masked and non-masked performances. Forming the difference signals yielded no consistent trends in masked/non-masked data.

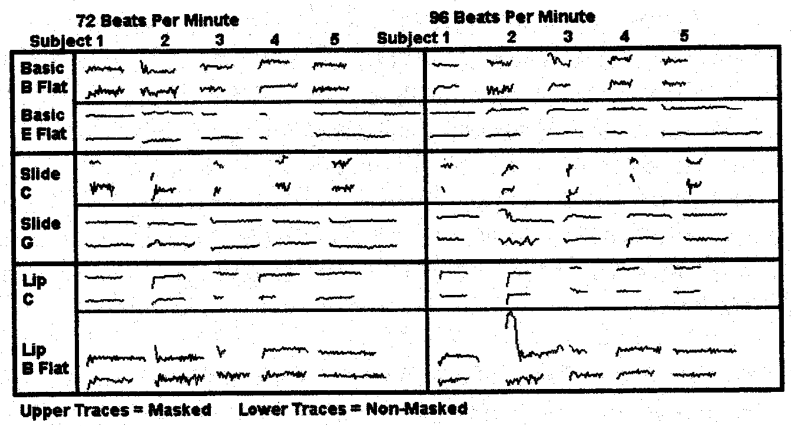

Figure 2 Log fundamental frequency extraction results from quasi-steady state tones played by all subjects. 72 and 96 bpm performances are shown. Masked performances are the upper traces in each pair.

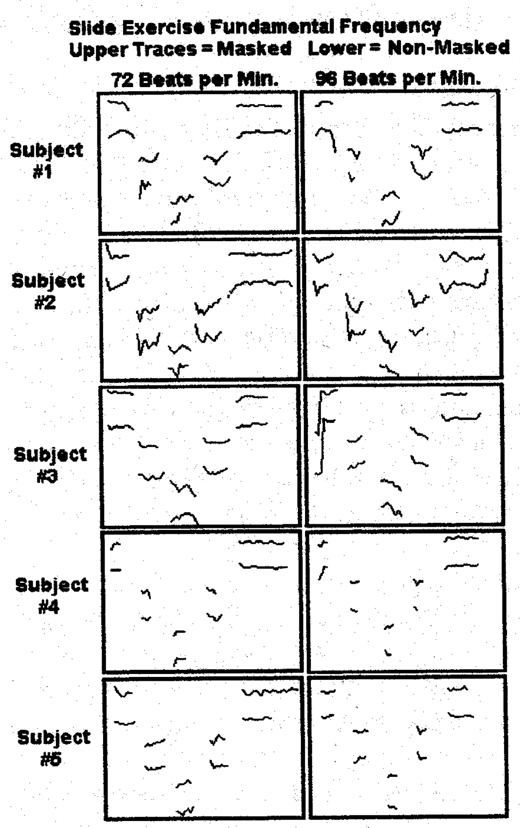

Figure 3 Slide exercise log fundamental frequencies from all subjects. Upper traces are masked performances.

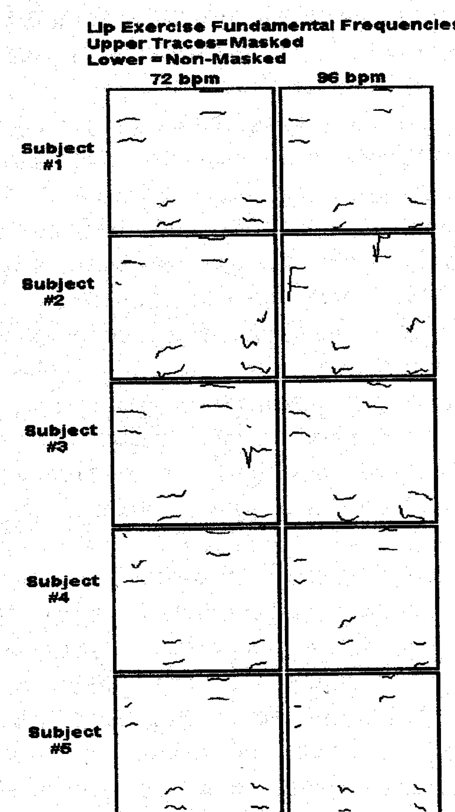

Figure 4 Lip exercise log fundamental frequencies from all subjects. Upper traces are masked performances.

Elliot & Niemoeller (1970) performed experiments in which singers matched target tones with and without auditory masking. They found that the first 1/2 second was unaffected by masking, but the remainder of the sung tone exhibited much pitch perturbation in the presence of auditory masking. This is consistent with other studies of the roles of memory versus feedback control channels, with memory playing the significant role before and during the first 250-500 ms., and feedback mechanisms dominating thereafter. To attempt to isolate the effects of memory from those of auditory and haptic feedback, average frequencies and standard deviations were calculated in the first 250 ms. of each cadence note, and compared to the remainder of the note. Again, little difference was found between non-masked and masked performances, except for the decline in standard deviation in the masked performances mentioned above.

Qualitative evaluations of the non-masked/masked pairs yielded many cases where, except for the increase in loudness, there is little difference between the two performances. Even though not asked, in most cases subjects announced which note(s) they had missed in a masked performance, and correctly identified which note they had played instead of the intended note. All players reported playing "by feel," and used terms to describe feeling the correctness of a pitch including "finding the notch," "centering the note," etc.

The results suggest that, for trombone players, memory plays a significant role in setting up the notes. Haptic channels are either dominant as the feedback channel, or suffice very well in the case of impaired auditory feedback. The memory hypothesis motivates further study using beginning players in which memory hasn't been significantly trained. The results also seem to support some brass literature which asserts that the principal mechanism for pitch correction in solo performance is feeling the 'center of the note.' This suggests further study in which the feedback to the lips is in some way altered.

Thanks always to the people at the Stanford Center for Computer Research in Music and Acoustics, where the experiments were performed. Thanks to Brent Gillespie and Dan Levitin for conversations both about the experiments and interpreting the data thereafter.

References

Arban, J. (1936), *Arban's Complete Conservatory Method for Trumpet*, E. Goldman and W. Smith eds., New York, Carl Fischer.

Askenfelt A. and E. Jansson (1992), On Vibration Sensation and Finger Touch in Stringed Instrument Playing, *Music Perception*, 9:3 pp. 311-350.

Brown, J. (1996), Frequency Ratios of Spectral Components of Musical Sounds, *Journal of the Acoustical Society of America*, 99:2, pp 1210-1218.

Campbell, W. and J. Michel (1980), The Effects of Auditory Masking on Vocal Vibrato, in V. Lawrence and B. Weinberg eds., *Transcripts of the Eighth Symposium on Care of the Professional Voice, Part I: Physical Factors in Voice, Vibrato, Registers.* New York: The Voice Foundation, pp. 50-56.

Deutsch, J. and J. Clarkson, (1959), Nature of the Vibrato and the Control Loop in Singing, *Nature*, 183: pp. 167-168.

Elliot, L. and A. Niemoeller (1970), The Role of Hearing in Controlling Voice Fundamental Frequency, *International Audiology*, 9: pp. 47-52.

Kleinhammer, E. (1963), *The Art of Trombone Playing*, Evanston Illinois, Summy-Birchard Company.

Lindqvist-Gauffin, J. and J. Sundberg (1974), Masking Effects of One's Own Voice, *Speech Transmission Lab Quarterly Research and Progress Report*, Royal Institute of Technology, Stockholm, 1:1974, pp. 35-41.

Møller, A. (1972), The Middle Ear, Chapter 4 in *Foundations of Modern Auditory Theory II*, J. Tobias ed., New York, Academic Press, pp. 134-194.

Royster, J. (1989), Sound Exposures and Hearing Thresholds of Musicians in a Major Symphony Orchestra, *Journal of the Acoustical Society of America*, Abstract, Vol. 85, Supplement 1, pp. S46.

Shipp, T., Sundberg, J. and S. Haglund, (1984), A Model of Frequency Vibrato, *13th Symposium on Care of the Professional Voice*, L. van Lawrence, ed. ,New York: Voice Foundation, pp. 116-117.

Sundberg, J. (1981), To Perceive One's Own Voice and Another Person's Voice, in *Research Aspects on Singing*, Stockholm, Royal Swedish Academy of Music, pp. 80-96.

Tonndorf, J. (1972), Bone Conduction, Chapter 5 in *Foundations of Modern Auditory Theory II*, J. Tobias ed., New York, Academic Press, pp. 197-237.

Ternström, S., Sundberg, J. and A. Colldén (1983), Articulatory Perturbation of Pitch in Singers Deprived of Auditory Feedback, *Speech Transmission Lab Quarterly Research and Progress Report*, Royal Institute of Technology, Stockholm, 2-3:1983, pp. 144-155.

Ward, D. and E. Burns (1978), Singing Without Auditory Feedback, *Journal of Research in Singing and Applied Vocal Pedagogy*, 1:2 pp. 24-44.