Auditory Discrimination for Frequency Modulation Extent in the Synthesized Vowel /α/

Yangho Cho
Korea Communication Disorders Research Institute

INTRODUCTION

The human phonatory mechanism has long been of special interest to medical scientists, speech scientists, speech pathologists, linguists, voice coaches and vocal performers. One of the intriguing phonatory maneuvers routinely produced by singers is vocal vibrato. In vocal vibrato, singers not only phonate at target frequencies above and below the target within about 0.5 to 2 semitones at approximately 4 to 8 times/sec. Thus, vocal vibrato exemplifies a complex frequency control maneuver.

The literature is equivocal on the issue of controllability of rate and/or extent of frequency modulation in vocal vibrato (1-5). A study, however, strongly indicated that the rate of frequency modulation is under voluntary control between 3 and 7 Hz while extent is not (6-7). On the average, singers matched rates of 3, 5, and 7 Hz within 10% and extent of o.5, 1 and 1.5 semitones within 60%. King's findings not only have significant implications to the theory of vocal vibrato production, but also naturally leads to another important area of investigation regarding frequency control capability.

It is reasonable to assume that one would not develop motor control unless the results of the control were perceptible. King's subjects demonstrated their ability to match the frequency modulation rate suggesting that they perceived the differences in the rate. One cannot conclude, however, that they did not perceive the differences in the extents simply because they could not match the extent. Possibly, they failed to control extent in spite of being able to hear differences in the extent.

The purpose of the present study was to examine listeners' ability to discriminate differences in magnitudes of extent of 5-Hz and 7-Hz frequency modulations. Knowledge of such perceptual constraints would provide further insight into human phonatory capabilities. Primary questions asked in this investigation were: (1) What are DLs defined as the 75% response level for the extent of frequency modulations in synthesized vibrato /α/? (2) Are the extent DLs different for 5-Hz versus 7-Hz modulation rates?

METHOD

Auditory Stimuli. Auditory stimuli were synthesized /ɑ/ with vocal vibrato generated by a VAX computer using the program SPEAK is a speech synthesis program developed by Titze (8).

Two sets of 12 synthesized /ɑ/s with differing amounts of extent of frequency modulation were used in this study. All stimuli had the base (fundamental) frequency modulation increased from 2% to 7.5% with a 0.5% increment. These percent values correspond to the extents of 0.68 semitones to 2.42 semitones with an increment of 0.16 semitones. The rate of frequency modulation for one set was 5 Hz and the other was 7 Hz. The duration of each stimulus was one second.

Auditory Discrimination Test Tapes. An Auditory discrimination test tape was created in a paired comparison paradigm for the 5-Hz and 7-Hz tests. A pilot study indicated that comparisons of two stimuli separated by more than six steps were clearly distinguishable. Therefore, it was decided that only stimulus pairs separated by six or less steps be used for the discrimination tests. Pairing in this manner resulted in a total of 51 possible pairs for each set.

Taking into consideration a possible order effect within each pair, it was decided to include, additionally, the same 51 pairs but to reverse the order of stimuli within each pair. The final discrimination test tape for each modulation-rate condition, therefore, contained a total of 102 pairs. The orders of the 102 pairs in the 5-Hz and 7-Hz tests were randomized.

Subjects and Experimental Procedures. A total of 50 adults, 29 males and 21 females, served as the subjects. Their age ranged from 22 to 46 with a mean of 31 years. The subjects passed a pure-tone hearing screening test at 500, 1000 and 2000 Hz at 20 dBHL in at least one ear. Of the 50 subjects, 9 were considered to be trained singers.

DL for the Extent of Frequency Modulation. As stated earlier, DL was defined as the 75% correct response level. As a preliminary to the DL calculation, for each subject, the total correctly responses in percent at each step difference were calculated for the 5-Hz and 7-Hz tests. The results are presented in Table 1 wherein the total number of responses for each modulation test, and means and standard deviations of percent correct responses, are shown for each of the step differences for the 5-Hz and 7-Hz

tests, respectively. As expected, correct responses increased as the step differences increased for both tests. For the 5-Hz test, the one-step difference yielded 66% correct responses, while two-step difference yielded 78% correct responses. DL being defined as the 75% correct response level, DL for this set of data was located between the one and two step difference (which corresponded to 0.5% and 1.0% frequency extent difference, respectively). DL was determined to be 0.85% frequency extent using interpolation.

For the 7-Hz test, the one step difference yielded 66% correct responses. The interpolation procedure yielded a DL of 0.81% frequency extent.

DISCUSSION

DL for the Extent of Frequency Modulation. The primary purpose of the present investigation was to find the DL for to extent of frequency modulations in vibrato and to gain insight why the singers in the King (6) study could not match extent of frequency modulation but could match rate of modulation. In particular, a question was raised whether the singers failed to match extent accurately (1) because they could not discriminate the extent, or (2) because they could discriminate but could not control extent in their phonatory attempts.

DL was found to be 0.29 semitones (peak-to-valley range) for the 5-Hz modulation stimuli, and 0.28 semitones for 7-Hz modulation stimuli. Production studies (9, 10, 11, 12, 13, 14, 15, 16), reported 0.5 semitones to 2.0 semitones as a typical range of frequency extent in vocal vibrato. The present results indicate that, in ideal conditions of sustained, steady vibrato vowels, one would be able to distinguish perceptually an extent difference of about 0.3 semitones.

Significant Step Difference Effects. Through a pilot study, it was expected that the DL would be located somewhere between o.5% (one-step difference) and 3.0% (six-step difference). Thus, it was not unexpected that the discrimination scores would result in an overwhelmingly significant step difference effect in the ANOVA test. What was not foreseen was the gradual, rather than sharp, improvement of the discrimination scores as the step difference increased. The discrimination scores reached a plateau only after the step difference was 2.5% (five-step difference) as evidenced by the post hoc Newman-Keuls test. The gradual, statistically significant, improvement of the discrimination scores up through the five-step difference appears to indicate that 1) the extent increment of 0.5% was adequate, 2) the

discrimination errors, although decreased, persisted even at the large step differences possibly due to relatively long signal durations for this kind of paired-comparison tasks.

The subject pool happened to include 9 music majors. Although the size of this subgroup was small and it was not in the original design, possible differences in the discrimination performances between the music and non-music majors were explored. The difference in the discrimination scores for both 5-Hz and 7-Hz tests between music major and non-music major subjects (176.85 vs. 170.12) was not significant at the 0.05 level ($t=0.49$, $df=48$).

CONCLUSION

1. The DLs calculated in this study were 0.29 semitones for the 5-Hz tests and 0.28 semitones for the 7-Hz tests.
2. The effects of auditory discrimination at the 5-Hz rate of frequency modulation in vocal vibrato are not different from those at 7-Hz.
3. The effects of step differences are significant.
4. There is no effect from the interaction of modulation rates and step differences.
5. There is no gender difference in frequency extent discrimination.
6. The differences in the discrimination scores for both 5-Hz and 7-Hz test between music major and non-music major subjects are not significant.

REFERENCES

1. Vennard, W. (1967). Singing: the mechanism and the Technique. Fisher, New York.

2. Ruhlman, J. (1984). Science Session III. In Lawrence, V. and Weinberg, B. (Eds.). Transcripts Thirteen Symp. Care Professional Voice, 118-129.

3. Shipp, T., Leanderson, R. and Sundberg, J. (1979). Vocal vibrato/rate and extent of vibrato as a function of vowel, effort and frequency. In Lawrence, V. and Weinberg, B. (Eds.). Transcripts Eighth Symp. Care Professional Voice, part I : 50-56.

4. Keidar, A., Titze, I. and Timberlake, c. (1984). Vibrato characteristics of tenors singing high C's. In Lawrence, B. and Weinberg, B. (Eds.). Transcripts Thirteen Symp. Care Profssional Voice, 105-110.

5. Campbell, W.M. and Michel, J.F. (1979). The effects of auditory masking on vocal vibrato. In Lawrence, V. and Weinberg B. (Eds.). Transcripts Eighth Symp. Care Professional Voice, 1:50-55.

6. King, J.B. (1989). Control of Rate and Extent of Frequency Modulation in Vocal Vibrato by Trained Singers, Doctoral Dissertation in Vocal Vibrato by Trained, Doctoral Dissertation. University of Colorado.

7. King, J.B. and Horii, Y. (1989). Rate and Extent Control of Vibrato by Trained Singers. ASHA, 31, 155.

8. Titze, I. (1983). Synthesis of sung vowels using a time-domain approach. Transcripts Eleenth Symp. Care of the Professional Voice, Part I: Scientific Sessions.

9. Gemelli, A., Sacerdote, G., and Bellusi, G. (1954). Analisi electroacustica della voce contat. Boll. Soc. Ital. Fonet Sper., 3-8.

10. Horii, Y. (1989a). Frequency modulation characteristics of sustained /a/ sung in vocal vibrato. Journal of Speech and Hearing Research, 32, 829-836.

11. Luchsinger, R., and Arnold, G. (1965). Voice-Speech-Language. Belm ont: Wadsworth.

12. Mason, R.M. (1965). A Study of the Physiological Mechanisms of Vocal Vibrato. Unpublished doctoral dissertation. Illinois: University of Illinois.

13. Mason, R.M. and Zemlin, W.R. (1966). The phenomenon of vocal vibrato. The NATS Bulletin, 22:12-17.

14. Seashore, C. (1938). The vibrato. Studies in the Psychology of Music, (Vol. 1), University of Iowa, Iowa City.

15. Sjostrom, L. (1947). Experimental-phonetischeutersunchugen des vibratophanomes der sinstimme. 10th Nord. Otolaryngol. Kongr. Stockholm.

16. Wincke, F. (1953). Physikalische kroterien fur objecktiv stimmbeurilung. Folia Phoniatrica, 5:232.

Table 1. Means and standard deviations of percent correct responses at each of the six step differences for the 5-Hz and 7-Hz tests.

Step Difference	Total No. of Responses	5-Hz % Correct		7-Hz % Correct	
		Means	SDs	Means	SDs
1	1100	66.2	10.0	65.7	10.4
2	1000	78.3	10.7	79.9	9.3
3	900	85.3	9.6	89.3	9.6
4	800	90.0	8.7	93.3	9.4
5	700	94.2	8.2	94.5	7.9
6	600	96.7	7.7	96.2	6.5