

Influence of Pitch Height on the Perception of Submissiveness and Threat in Musical Passages

DAVID HURON

The Ohio State University

DARYL KINNEY

The Ohio State University

KRISTIN PRECODA

SRI International

ABSTRACT: Bolinger, Ohala, Morton and others have established that vocal pitch height is perceived to be associated with social signals of dominance and submissiveness: higher vocal pitch is associated with submissiveness, whereas lower vocal pitch is associated with social dominance. An experiment was carried out to test this relationship in the perception of non-vocal melodies. Results show a parallel situation in music: higher-pitched melodies sound more submissive (less threatening) than lower-pitched melodies.

Submitted 2006 April 6; accepted 2006 April 14.

KEYWORDS: *music, prosody, pitch height, emotion, threat, aggression, timidity, politeness*

RESEARCH by Bolinger, Ohala, Morton and others has established that vocal pitch height (F0) is perceived to be associated with social cues for dominance and submissiveness. In a wide sample of cultures, Bolinger (1964) noted that high or rising vocal pitch is associated with politeness, deference, submissiveness and lack of confidence. Bolinger also showed that conversely, low or falling vocal pitch is associated with authority, threat, aggression, and confidence. Ohala (1984) and Morton (1994) have assembled additional support for this association in ethological studies of non-human animals.

Paradoxically, an earlier study by Scherer *et al.* (1973) seemed to suggest that high vocal pitch is associated with aggression; however, Ohala (1994) has noted that the materials studied by Scherer *et al.* exhibit marked descending pitch contours. That is, beginning from a high pitch may allow a more exaggerated pitch descent, corresponding to a presumed confident or aggressive assertion. This association has been demonstrated in both ecological and controlled experimental situations.

Morton (1994) reviewed vocalizations for 54 species as observed by ethologists. Morton noted that three features appear to be important in aggressive/passive signaling. First, low-pitched sounds are generally associated with aggressive signaling, whereas high-pitched sounds are generally associated with friendly, appeasing, or fearful signals. Second, falling pitch contours are generally associated with aggression, while rising pitches are associated with friendly, appeasing, or fearful signals. Finally, Morton drew attention to the (periodic) tone versus (aperiodic) noise distinction. In general, hostile signals are described as raspy, growling, buzzing, or snarling. Such sounds are often characterized as "harsh." By contrast, friendly or appeasing signals are typically described as whimpers, whines, squeaks, squeals, screeches, or chirps. In general, those tones which evoke the clearest pitch sensations are associated with friendly and appeasing signals, whereas unpitched noises or less clearly pitched tones are associated with aggression.

To Morton's three factors we may add a fourth – loudness. Loudness seems to be interpretable in two ways. First, loudness may be associated with the *urgency* of the signal. That is, a loud vocalization may indicate a strong desire to communicate or to communicate clearly. In addition, increased loudness is also likely correlated with hostility or aggression; the acoustic power might suggest the physical power of the individual or signal the individual's willingness to engage in physical confrontation.

Morton (1994) presented a model of species vocalizations consisting of two independent dimensions: fear and aggression. In agreement with Ohala, Morton has argued that aggression is associated with efforts to appear large.

Musical Repercussions

This research in vocal prosody suggests a number of possible parallels with musical organization and the ensuing musical semantics. If the acoustic connotations identified by Bolinger, Ohala, Morton and others generalize beyond speech and non-human animal vocalizations, then we might expect to observe similar effects in vocal and non-vocal music. That is, parallel connotations would be expected from the effects of pitch height, pitch contour, loudness, and toneness (i.e., the subjective correlate of periodicity).

Specifically, if Ohala's theory of the relationship between pitch height and dominance/submissiveness is correct, then one might predict that the semantic associations evoked by a given musical passage might change depending on transposition. In addition, it is plausible that falling and rising pitch contours might evoke assertive and tentative connotations, respectively. Similarly, noisy timbres might be evaluated by listeners as sounding more aggressive than more tonal timbres. Finally, louder sounds might be rated as either more confident or more urgent than softer sounds. In the case of music, a first step would be to determine whether the correlation between dominance/submissiveness and pitch height found by Ohala and others holds for musical stimuli.

Hypothesis

In light of the extant research, we proposed to test the hypothesis that transpositions of melodies have dominant/submissive connotations in proportion to the amount and direction of pitch transposition. We predict that versions of a melody will be perceived as sounding more aggressive, threatening, or dominant when transposed down in pitch, and more passive, submissive, or polite when transposed up in pitch.

METHOD

In order to test this hypothesis, an experiment was carried out wherein a set of 36 melodies was played to 41 listeners who rated the passages on six semantic scales.

Stimuli

The stimuli consisted of 36 phrases selected from a variety of melodies from several styles and cultures. The phrases were extracted from 3 Schubert Lieder, 3 Weber songs, 18 traditional European folk songs (Czech, Danish, Dutch, English, Hungarian, Romanian, Swiss), 2 Hassidic sacred songs, 1 Gregorian chant, 3 traditional African songs (Pondo, Xhosa, Zulu), 2 traditional Chinese songs, 1 Ojibway song, 1 American popular song, and 2 randomly generated tone sequences. An effort was made to select materials that the participants were unlikely to have heard.

Of those passages exhibiting a clear tonality, 15 were in major keys and 10 were in minor keys; 8 passages were modal, while 3 passages were in atonal or indeterminate modes. In order to avoid possible tonic effects, the original passages were transposed by modest amounts to ensure that no particular key (such as C major) predominated. During the experiment, each phrase was presented three times, with each presentation at a different pitch height: (1) normal, (2) transposed upward by 12 semitones, and (3) transposed downward by 12 semitones. In total, 108 stimuli were presented (36 phrases x 3 transpositions). In order to control for, or identify, possible timbre effects, subjects were assigned to one of two groups. The stimuli for one group consisted of sine tones, while the stimuli for the second group consisted of synthesized piano tones. Stimuli were presented in unique random orders for each participant with the constraint that no phrase be repeated within a window of eight stimuli.

Participants

The participants were 41 music students recruited from a second-year aural skills course at the Ohio State University. This experiment was one of several in which students could choose to participate for course credit. Of the 41 participants, 14 were male and 20 were female; the gender of 7 participants was not recorded.

Procedure

Each participant was tested separately. Melodic phrases were presented via stereo headphones and sounds were generated via a MIDI synthesizer. For each trial, subjects judged the quality of the stimulus using a 7-point Likert scale.

In any task involving semantic judgments, possible confounds can arise due to the particular terms used in the experiment (see e.g. Van de Geer, Levelt and Plomp, 1962). Consequently, several synonyms were used for the terms submissive and dominant. These included "politeness" and "timidity" as synonyms for submissiveness; and "aggressivity," "threateningness," and "heaviness" as synonyms for dominance. The term "heaviness" was also included in order to more directly address Ohala's claim that sounds are used to infer size, and that apparent size is the principal cue through which submission or dominance are communicated.

Each participant judged all stimuli according to just one term. Since our subjects were musicians, we elected to avoid having them judge "dominance" due to a fear that the term would evoke harmonic implications (i.e., "sounding in the dominant"). In total, six terms were used in the experiment: aggressive, threatening, heavy, polite, timid, and submissive.

Instructions

Participants received the following instructions:

In this experiment, you will hear a number of melodic passages. The passages will differ in tempo, transposition, meter, key, major/minor mode, style, phrasing, and other characteristics.

We would like you to characterize the quality of these passages. Specifically, we would like you to judge how _____ the passage is. This may seem like an odd way to judge musical passages, but please try to do your best anyway.

After each passage, you will be prompted to enter a numerical judgment. Enter a number between 1 and 7. Enter the number 7 if you judge the passage to be very _____. Enter the number 1 if you judge the passage to be very UN-_____. The value 4 lies in the middle of the range. Please try to use the entire scale (from 1 to 7) in your responses.

After typing your response, press the ENTER key to continue with the next passage.

RESULTS

Figures 1 and 2 show the mean ratings for the three transposition levels across each of the six semantic judgments. Figure 1 shows the results for aggressivity, heaviness, and threateningness; Figure 2 shows the results for submissiveness, timidity, and politeness. In Figure 1, notable declines are evident for heaviness and threateningness, as predicted. A very weak (non-significant) decline is also evident for aggressivity. In Figure 2, submissiveness and politeness show a predicted increase with transposition level; however, timidity judgments exhibit a V-shaped result, not consistent with the hypothesis. Note that the effect size for heaviness is extremely marked compared with the other judgments, showing a decline in mean ratings from nearly 6 to just over 2 on a 7-point rating scale. (The Y-axis range is larger in Figure 1 than in Figure 2.)

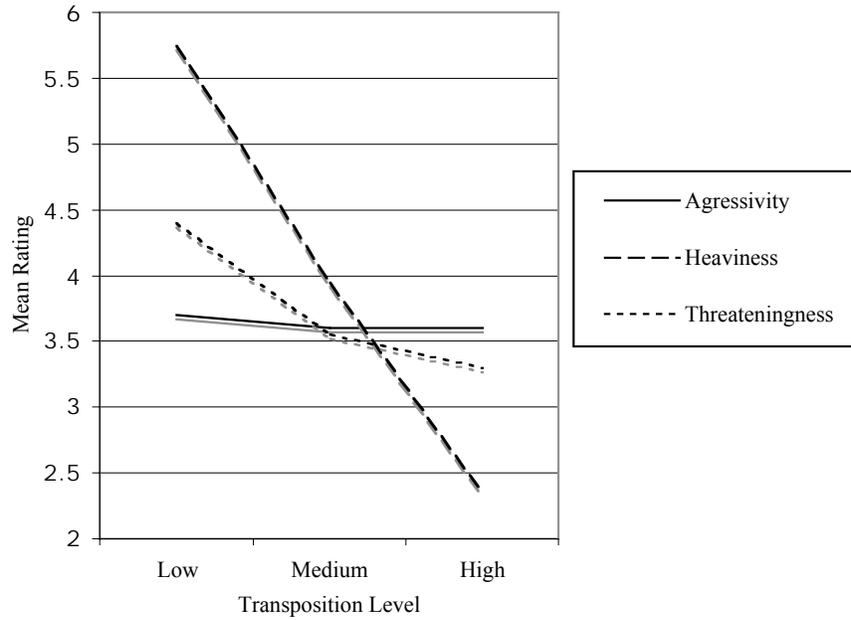


Fig. 1. Effect of pitch height for 36 phrases on judgments of aggressivity, heaviness, and threateningness. Lower transpositions evoke significantly increased judgments for heaviness and threateningness. No effect is evident for judgments of aggressivity.

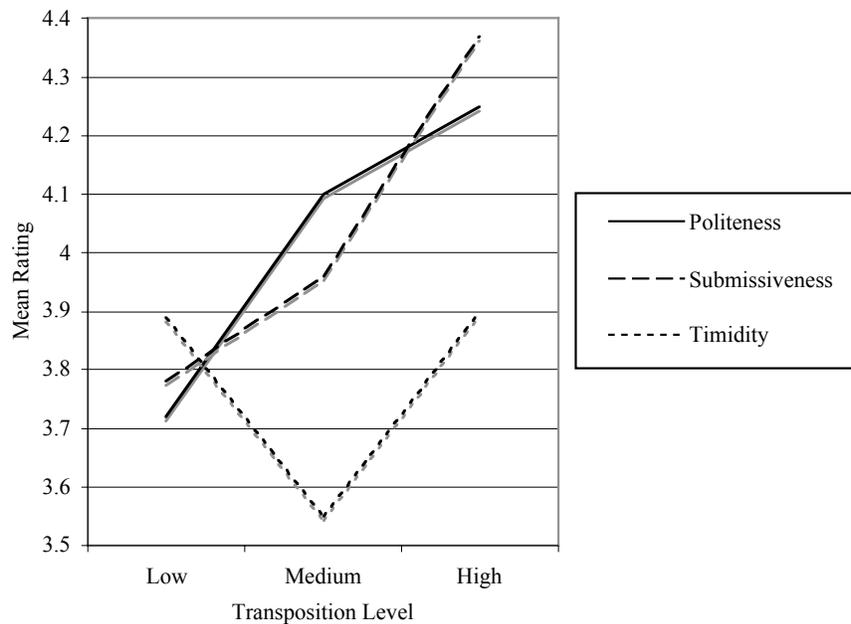


Fig. 2. Effect of pitch height for 36 phrases on judgments of politeness, submissiveness, and timidity. Higher transpositions evoke judgments of significantly greater politeness and submissiveness. Timidity exhibits a V-shaped response. An interaction was found (not shown) between timidity and timbre: for the piano timbre, an increase in pitch is associated with a significant increase in timidity, while for sine tones, an increase in pitch is associated with a significant decrease in timidity (see text).

Statistical models were fit to the data using logistic regression techniques. The goal of logistic regression is to predict the distribution of an ordered categorical response variable, in this case, ratings on a scale from 1 to 7. Variables used as predictors may be either categorical (e.g. melody identity, timbre, subject identity) or continuous (e.g., pitch). The estimated coefficients of the predictor variables may be examined for significance, and thus the predictor variables can be considered to play a significant or insignificant role in fitting the model to the observed distribution of the ratings.

Models were fitted for each semantic term using the LOGISTIC procedure in SAS and the default (logit) link. While not all the possible predictor variables could be used in all models, because of data sparsity, the mathematical expressions are of the form of the equations below. It should be noted, in particular, that since each subject heard either piano or sine stimuli but not both, and the number of subjects is small, it is questionable whether we can distinguish differences due to timbre and differences due to subject factors alone. In the analyses, models were based on timbre where possible, and on subject where the stimuli did not differ in timbre (i.e., for politeness and threateningness). It is possible, however, that some of the reported timbre differences might in fact be subject differences. In the equations below, *pitch* is the average pitch of the stimulus in semitones relative to middle C, *timbre* equals 1 for piano stimuli and -1 for sine tones, and each melody and subject has its/her/his own coefficient.

$$\begin{aligned} \gamma_j &= Pr(\text{rating} \leq j), \text{ the probability that a response is less than or equal to } j, 1 \leq j \leq 7 \\ \pi_j &= \gamma_j - \gamma_{j-1} = Pr(\text{rating} = j), \text{ the probability of response } j, 1 \leq j \leq 7 \\ \log \left(\frac{\gamma_j}{1 - \gamma_j} \right) &= \theta_j + \beta_{pitch} \cdot \text{pitch} + \beta_{timbre} \cdot \text{timbre} + \beta_{melody} + \beta_{subject} + \beta_{timbre \times \text{pitch}} \cdot \text{timbre} \cdot \text{pitch} \end{aligned}$$

The θ_j are analogous to the intercept term in a standard linear regression model, and affect the predicted distribution of ratings independently of any of the predictor variables. Because of this independence, they are of no further interest here. Most directly interpretable in the above equations are the π_j , probability that a rating is j , and the γ_j , probability that a rating is j or smaller. Because of the form of the left-hand side of the last equation above, the coefficients of the predictor variables are not as easily interpretable. However, it can be said that larger coefficients lead to a prediction of lower ratings, while smaller or more negative coefficients lead to a distribution more heavily weighted toward higher ratings. The estimated coefficients of the predictor variables and their p -values are summarized in Table 1.

Table 1. Coefficients and p -values of predictor variables. Cells are labeled *ne* for coefficients that could not be estimated from the data. Significant effects are shown in boldface.

term	Coefficient of						
	pitch	timbre	interaction of timbre and pitch	least melody	most melody	least subject	most subject
Aggressivity	-0.00146 <i>p</i> = .83	0.4981 <i>p</i> < .0001	0.00735 <i>p</i> = .25	1.85	-3.80	<i>ne</i>	<i>ne</i>
				<i>p</i> < .0001			
Heaviness	0.2671 <i>p</i> < .0001	0.4088 <i>p</i> < .0001	-0.0354 <i>p</i> < .0001	2.83	-1.85	<i>ne</i>	<i>ne</i>
				<i>p</i> < .0001			
Politeness	-0.0267 <i>p</i> < .0001	<i>ne</i>	<i>ne</i>	3.22	-3.58	0.31	-0.45
				<i>p</i> < .0001		<i>p</i> = .0108	
Submissiveness	-0.0282 <i>p</i> < .0001	0.1369 <i>p</i> = .07	-0.00727 <i>p</i> = .22	2.39	-1.15	<i>ne</i>	<i>ne</i>
				<i>p</i> < .0001			
Threateningness	0.0686 <i>p</i> < .0001	<i>ne</i>	<i>ne</i>	3.29	-3.21	0.38	-0.70
				<i>p</i> < .0001		<i>p</i> = .0003	
Timidity	0.00595 <i>p</i> = .37	0.1731 <i>p</i> = .02	-0.0391 <i>p</i> < .0001	2.28	-1.59	<i>ne</i>	<i>ne</i>
				<i>p</i> < .0001			

The results can be summarized as follows: increasing pitch does lead to judgments of significantly greater politeness and submissiveness; it also leads to judgments of decreasing heaviness and threateningness. There is no support for a significant effect of pitch on judgments of aggressivity or timidity. However, for timidity, pitch interacts significantly with timbre: for the piano, an increase in pitch is associated with a significant increase in timidity, while for sine tones, an increase in pitch is associated with a significant decrease in timidity.

Of course different melodies are significantly associated with different judgments for all six qualities judged. That is, some melodies were judged more aggressive or more polite than other melodies, no matter what the pitch register. The differences between the two most extreme melodies for each quality are greater than the difference due to any other factor. In other words, there exist one or more other factors in melodic organization that have a greater influence on perceived heaviness/aggressivity/threateningness/politeness/submissiveness/timidity.

With regard to the effect of timbre, for heaviness and timidity judgments, pitch height has a significantly greater effect for sine stimuli than for piano tones. For aggressivity and submissiveness no interaction was found between pitch and timbre. Possible interactions for politeness and threateningness could not be tested due to insufficient data.

Reported Strategies

In post-test interviews we asked our subjects to describe what strategies they were using in making their judgments. Nine subjects reported that their judgments were made predominantly from intuition and offered no comment on their strategies. For example, one subject said that she imagined meeting people in a hallway, and that the melodies "were what they were saying." Those subjects who spoke in more detail identified a variety of factors influencing their judgments including degree of chromaticism, major/minor mode, note durations, tempos, rhythms, tone color, tonal/atonal, and style. One subject reported answering predominantly based on whether he liked or disliked the melody.

Eleven subjects noted that whether a melody was major or minor influenced their responses. Eight subjects reported that tempo or duration was influential, with slower tempos typically sounding less aggressive. Without prompting, twelve subjects explicitly noted that pitch height was a factor influencing their judgments. Tellingly, one subject (who judged heaviness) reported that she focused primarily on duration and explicitly discounted pitch height, saying, "People would obviously tend to think of low pitches as heavier."

DISCUSSION AND CONCLUSION

The results are broadly consistent with the experimental hypothesis. Judgments of aggressivity and timidity did not show the expected effect. However, assembling all of the synonyms together, the results are clearly consistent with the views of Bolinger, Morton and Ohala concerning the relationship between pitch height and dominance/submissiveness. Ohala in particular has argued that the origin of this relationship may be found in the general correspondence of pitch height with size of the sound resonator. Both Ohala (1984) and Morton (1994) have suggested that perceptions of submissiveness and aggression in animals originate in displays of physical size (deference and covering are signaled by efforts to appear small; threat and aggression are signaled by efforts to appear large).

Interestingly, Ohala has extended this claim to account for the origins of the human smile. In most animals, displaying one's teeth is associated with a threat display. Paradoxically, this association does not apply in the case of primates. In all primates, displaying one's teeth is typically construed as a sign of submission rather than threat. In humans, smiling often results in the showing of teeth. Ohala has argued that the smile originated as an auditory display and only later became a visual display. The flexing of the zygomatic muscles of the checks causes a reduction in the size of the mouth cavity with a consequent shifting upward of the vocal formant frequencies. According to Ohala, this upward shift evokes an association with smaller size and postural deference that is common throughout the animal kingdom.

In controlling for timbre, we expected no significant difference between piano tones and sine tones. However, significant timbre effects were found for timidity, submissiveness, and heaviness. For submissiveness, the magnitude of the effect was found to be greater for piano tones, whereas for heaviness, the magnitude of the effect was found to be greater for sine tones. In the case of timidity, an increase in

pitch is associated with a significant increase in timidity for piano tones, while an increase in pitch is associated with a significant decrease in timidity for sine tones.

Of the six terms investigated in this study, the pitch effect size was greatest for judgments of heaviness. Recall Ohala's original view was that pitch is related to size, and that it is the size-related cues that lead to perceptions related to dominance and submissiveness. Recall also that Scherer *et al.* (1973) paradoxically found that higher vocal pitch can be associated with aggression. It may be that in some circumstances higher pitches are associated with a scream. Indeed, in post-test interviews with our subjects, four subjects made comments related to this idea. One subject reported that "the most aggressive sounds were either high or low; less aggressive in the mid-range." A second subject reported that she heard some high-pitched melodies as being more threatening. A third subject reported that high sounds were "sometimes whiny," and a fourth reported that high pitches were sometimes piercing and thus less timid.

It is also possible that pitch height interacts with loudness. Screams or shouts are typically loud compared with submissive squeaks or whimpers. Moreover, the unpredicted results for timidity and aggressivity may relate to potency, as signaled through increased amplitude. Future research along these lines might explore the hypothesis that timidity and aggressivity are predominantly associated with loudness, with a possible small interaction with pitch. One can be threatening without necessarily being loud. One can be polite without necessarily being quiet. But it is more difficult to imagine being actively aggressive in the absence of loudness, and timid in the absence of quiescence.

Overall, this experiment is consistent with research in speech showing an association between pitch height and perceptions of submissiveness/dominance. Our research further establishes that this relationship generalizes to auditory stimuli beyond speech and animal vocalizations. Specifically, transposing the pitch height for musical passages has predictable effects on listener perceptions akin to social signals for dominance and submissiveness.

Informal observations of the music literature appear to echo the relationship between pitch height and social relationship found in this experiment. Composers may well intentionally place a melody in a lower register in order to evoke threatening, dominant or aggressive associations. Conversely, the transposition of a melody upward may be done to evoke more passive, vulnerable or submissive associations. Treatises on orchestration such as those of Rimsky-Korsakov (1891; e.g. pp.19, 24), Keenan and Grantham (1990; e.g. pp.77, 95, 99) and Piston (1955; e.g. p.168) explicitly describe lower registers as heavy, dark, and menacing, and higher registers as soft, light, and gracious.

REFERENCES

- Apple, W., Streeter, L.A., & Krauss, R.M. (1979). Effects of pitch and speech rate on personal attributions. *Journal of Personality and Social Psychology*, Vol. 37, pp. 715-27.
- Bauer, H.R. (1987). Frequency code: Orofacial correlates of fundamental frequency. *Phonetica*, Vol. 44, pp. 173-91.
- Bolinger, D.L. (1964). Intonation across languages. In: J.H. Greenberg, C.A. Ferguson & E.A. Moravcsik (Eds.), *Universals of Human Language, Vol. 2: Phonology*. Stanford, CA: Stanford University Press, pp. 471-524.
- Brown, B.L., Strong, W.J., & Rencher, A.C. (1974). Fifty-four voices from two: The effects of simultaneous manipulations of rate, mean fundamental frequency, and variance of fundamental frequency on ratings of personality from speech. *Journal of the Acoustical Society of America*, Vol. 56 No. 2, pp. 313-8.
- Cruttenden, A. (1981). Falls and rises: Meanings and universals. *Journal of Linguistics*, Vol. 17, pp. 77-91.
- Fairbanks, G., & Pronovost, W. (1939). An experimental study of the pitch characteristics of the voice during the expression of emotion. *Speech Monographs*, Vol. 6, pp. 87-104.
- Kennan, K., & Grantham, D. (1990). *The Technique of Orchestration*. Englewood Cliffs, NJ: Prentice-Hall, 4th edition.

- Van de Geer, J.P., Levelt, W.J.M. & Plomp, R. (1962). The connotation of musical consonance. *Acta Psychologica*, Vol. 20, pp. 308-19.
- Morton, E. (1994). Sound symbolism and its role in non-human vertebrate communication. In L. Hinton, J. Nichols & J. Ohala (eds.), *Sound Symbolism*, Cambridge: Cambridge University Press, pp. 348-65.
- Ohala, J. (1978). The production of tone. In V.A. Fromkin (ed.), *Tone: A Linguistic Survey*. New York: Academic Press, pp. 5-39.
- Ohala, J. (1980). The acoustic origin of the smile. *Journal of the Acoustical Society of America*, Vol. 68, p. S33.
- Ohala, J. (1982). The voice of dominance. *Journal of the Acoustical Society of America*, Vol. 72, p. S66.
- Ohala, J. (1983). Cross-language use of pitch: an ethological view. *Phonetica*, Vol. 40, pp. 1-18.
- Ohala, J. (1984). An ethological perspective on common cross-language utilization of F0 in voice. *Phonetica*, Vol. 41, pp. 1-16.
- Ohala, J. (1994). The frequency code underlies the sound-symbolic use of voice pitch. In L. Hinton, J. Nichols & J. Ohala (eds.), *Sound Symbolism*, Cambridge: Cambridge University Press, pp. 325-47.
- Piston, W. (1955). *Orchestration*. New York: W.W. Norton.
- Rimsky-Korsakov, N. (1891/1922/1964) *Osnovy orkestrivki*. Moscow, 1891. Berlin: Edition Russe de Musique, 1922. Translated by Edward Agate as *Principles of Orchestration*. Reprinted New York: Dover Publications, 1964.
- Scherer, K., London, H., & Wolf, J.J. (1973). The voice of confidence: Paralinguistic cues and audience evaluation. *Journal of Research in Personality*, Vol. 7, pp. 31-44.
- Tartter, V.C. (1980). Happy talk: Perceptual and acoustic effects of smiling on speech. *Perception & Psychophysics*, Vol. 27 No. 1, pp. 24-7.
- Williams, C.E., & Stevens, K.N. (1972). Emotions and speech: Some acoustical correlates. *Journal of the Acoustical Society of America*, Vol. 52 No. 4, pp. 1238-50.