Eyebrow movements and vocal pitch height: Evidence consistent with an ethological signal

David Huron[a] and Daniel Shanahan[b]
Cognitive and Systematic Musicology Lab, School of Music, Ohio State University, 110 Weigel Hall, Columbus, Ohio 43210

(Received 3 December 2012; revised 26 February 2013; accepted 18 March 2013)

When asked to sing a high pitch, people produce a facial expression that is judged more friendly compared with singing a low pitch [Huron *et al.* (**2009**). Empirical Musicology Rev. **4**(3), 93–100]. This effect was observed even when judges viewed only the face above the tip of the nose, and implies a relationship between pitch height and eyebrow height. In the current study, we examine the reverse relationship. Thirty-one participants were asked to read aloud standard texts while holding their eyebrows in a raised, neutral, or lowered position. Average $F0$ was found to correlate positively with eyebrow position, with higher vocal pitch associated with higher eyebrow placement. However, manipulating eyebrow placement produces a considerably smaller effect (on pitch) compared with the effect of manipulating pitch (on eyebrows). Results are discussed from the perspective of ethological signals [Lorenz (**1939**). Zool. Anz. **12**, 69–102].
 [http://dx.doi.org/10.1121/1.4798801]

PACS number(s): 43.70.Ep, 43.70.Aj, 43.66.Ba, 43.70.Bk [DD] Pages: 2947–2952

I. INTRODUCTION

In animal behavior (ethology) a useful distinction is made between a "signal" and a "cue" (e.g., Smith and Harper, 2003). A signal is a functional communicative act involving innate behavioral and physiological mechanisms (Lorenz, 1939; Hasson, 1994). A good example of a signal is a rattlesnake's rattle. The rattle is used as a warning in situations where the snake experiences fear. The rattle communicates to the fear-inducing animal that the snake could retaliate or respond in a life-threatening way. By contrast, a cue is an inadvertent conveyance of information that is artifactual and unintended. For example, in order for a snake to attack, it must adopt a coiled posture. If the snake is fully stretched out, it cannot strike. If we observe that a snake assumes a recoiled posture, we might rightly interpret this as a prelude to attack. But unlike the shaking of the rattlesnake's rattle, the pre-strike posture is an artifact of the mechanics required to attack. Although both the rattle and the posture are informative to an observing animal, the rattle behavior is regarded as a signal, whereas the recoiled posture is regarded as a cue.

Ethologists have noted that one of the ways in which signals can be distinguished from cues is that signals tend to exhibit *redundancy* where the signal is repeated or sustained over time and over multiple channels (Wiley, 1983; Johnstone, 1997). Since signals are intended to be communicated, signals should be obvious, rather than subtle. Employing more than one sensory modality makes the signal more conspicuous. For example, in the case of the rattlesnake's rattle, there is both a distinctive acoustical component (the sound of the rattle) as well as a distinctive visual component (the raised shaking tail). Ostensibly, even if an observer is only able to hear, or just see the snake, the signal could nevertheless be successfully communicated. By contrast, many (though not all) cues do not exhibit multimodal features. This simply reflects the fact that cues are artifacts, and not explicitly intended to be communicative.

In general, psychological research on emotion has been slow to recognize the value of ethological concepts in understanding emotional displays. For example, most of the research on the human smile has focused on the visual elements involved in smiling. The linguist John Ohala (1980, 1982) however noted that smiling has a distinctive effect on the voice. Apart from seeing a smile, you can also hear a smile (Tartter, 1980). Specifically, flexing the zygomatic muscles causes the lips to be pulled taut against the teeth, shortening the effective vocal tract length (VTL) of the speaker. This contraction produces an upward shift in vocal resonance that is easily heard as "smiling voice." Conversely, when the zygomatic muscles are relaxed, the lips tend to thicken and slump away slightly from the teeth, lengthening the VTL and producing a distinctively "darker" timbre associated with sadness or seriousness (Ohala, 1984).[1] Although smiling may involve closed lips with no sound production, in the social contexts in which smiling occurs there is a strong likelihood of vocal interaction. Ohala has suggested that the open-mouthed smile with accompanying vocalization may represent the original display behavior, with the closed-mouth version arising later through ritualization.

In an earlier study (Huron *et al.*, 2009), we asked participants to sing moderate, high, and low pitches while their faces were photographed. In a two-alternative forced choice task, independent judges selected the high-pitched faces as "more friendly" than the low-pitched faces. The same result was obtained when photographs were cropped to reveal only

[a]Current address: School of Music, 1866 College Road, Columbus, OH 43210.
[b]Author to whom correspondence should be addressed. Electronic mail: shanahan.37@osu.edu

the region above the tip of the nose. That is, judges still rated the faces associated with producing a high pitch as friendlier than those associated with a low pitch—even though the photograph omitted the lower region of the face. Although no formal analyses of the facial features were carried out, it was apparent that the eyebrows seemed to provide the most salient cue. Eyebrows tended to be raised for high pitches, and lowered for low pitches, at least in the case of Western-enculturated participants.

This pattern has also been observed in more ecologically valid musical contexts. For example, Bonfiglioli *et al.* (2006) examined the facial expressions of famous pianists performing on video recordings. They then related the eyebrow movements to musical features, such as dynamics, articulation, and pitch register. Based on descriptive non-quantitative observations, they reported that the players tend to raise their eyebrows when playing in medium and high pitch registers especially when performing at quiet dynamic levels. In addition, they observed that pianists tend to frown when performing at medium and low pitch registers at louder dynamic levels.

Ethologist Eugene Morton reviewed the vocalization repertoires for 58 species, including mammals and birds. Morton found a cross-species pattern in which high frequencies are associated with affiliative behaviors, whereas low frequencies are associated with aggressive behaviors (Morton, 1977, 1994). In a cross-cultural study of speech prosody, Bolinger (1964) found a similar association in humans: A high $F0$ is associated with friendly behaviors whereas a low $F0$ is associated with aggressive behaviors. Ohala (1994) has suggested that the origin of these associations arise from a basic acoustical fact: One of the best generalizations one can make about sound is that low frequencies are produced by large masses or large resonant cavities, whereas high frequencies are produced by small masses or small resonant cavities. Ohala invoked "sound-size symbolism," in which large (low) and small (high) frequencies parallel visual displays in which appearing large connotes aggression whereas appearing small connotes appeasement or affiliation.

Ohala argued that the "smile" is a combination of distinctive visual and distinctive acoustic elements. To Ohala's observations, we might add that this cross-modal pattern is consistent with the *redundancy* criterion used by ethologists to help distinguish signals from cues. By exhibiting a distinctive sound, the smile can be detected even when an observer is unable to see a person's face; by exhibiting a distinctive visual feature, the smile can be detected even when no sound is produced or heard. Once again, this multi-modal aspect reinforces the view that the human smile exhibits the hallmarks of an ethological signal.

Apart from the smile, research on facial expressions has similarly documented other characteristic features associated with friendliness and aggression. For example, lowered eyebrows are known to be an important component of aggression displays (Ekman and Friesen, 2003). Given the cross-modal association observed in Huron *et al.* (2009) one might conjecture that vocal pitch and eyebrow placement represent two components of a single signal system. If such a signal system exists, then one might expect a single neural source, which, when activated, would account for both the motor movements controlling eyebrow position as well as the tension of the vocal folds. That is, those movements involved in changing the pitch of the voice and those controlling the eyebrows would be conjectured to have a common source. In Huron *et al.* (2009), participants were asked to sing high and low pitches. At face value, the results imply that participants executed voluntary motor movements intended to manipulate the voice, and that eyebrow movements were observed as an incidental artifact. Another possibility is that, when asked to lower the pitch of the voice, participants unconsciously coded this request as one to produce an aggressive display, and therefore the aggression signaling system was activated with the consequence of the eyebrows moving as well. If this latter interpretation is correct, then we should observe a similar pattern of behavior when participants are instructed to move their eyebrows, rather than change the pitch of their voice. That is, we might theorize that both vocal pitch movements and eyebrow movements are mediated by a common motor circuit related to specific affective displays. At the same time, there is at least one other hypothesis that might predict the same observations: What might be dubbed the Hebb's Rule Hypothesis ("cells that fire together, wire together"). In the same way that scratching a cat's ear may cause the cat's leg to move in a scratching-like manner, the frequent coordination of vocal pitch and facial expression may produce a neurological association. We will address these questions later in Sec. VII.

II. HYPOTHESIS

In formal terms, we may state our hypothesis as follows:

H1. Raising one's eyebrows will tend to raise the pitch of the voice.

Conversely, lowering one's eyebrows will tend to lower the pitch of the voice.

III. SUBJECTS

For this experiment, a convenience sample of 31 subjects was recruited. Participants were undergraduate music students from the Ohio State University School of Music Subject Pool who received partial course credit for their involvement. In total, 17 female and 14 male subjects participated.

IV. PROCEDURE

In brief, participants read aloud short texts with their eyebrows placed in a neutral, low, or high position. Sound recordings were made of the spoken texts and the average fundamental pitch was determined for each trial.

Participants were tested individually in an Industrial Acoustics Corporation sound-attenuated room. Participants sat in front of a computer monitor that randomly displayed one of 12 English sentences as well as icons indicating how they should place their eyebrows. In many cultures, terms cognate with "high" and "low" are used to describe pitch. That is, pitch is commonly described using a height metaphor. Since we are instructing participants to move their eyebrows "up" or "down" there may be some unintended cross-modal spill-over due to the common metaphor. Consequently, we endeavored to create a procedure in which we could instruct participants to move their eyebrows up or down, without employing a height metaphor. Specifically, we avoided the use of vertical arrows, verbal instructions involving vertical connotations, or the vertical placement of material on the screen.

Instead, we made use of two icons, placed horizontally. One icon was a clip-art image of a nose; the second icon was a clip-art image of a hair-line. In instructing participants, we told them to move their eyebrows either "toward your nose" or "toward your hairline." Specifically, in each trial a horizontal arrow pointed to one of the two clip-art images. In order to ensure compliance with the eyebrow placement instructions, the experimenter remained with the participant throughout the experiment.

In order to reduce demand characteristics, participants were told that the purpose of the experiment pertains to auditory memory, and that the eyebrow placement was employed to increase the task difficulty.

V. INSTRUCTIONS

Participants received the following instructions:

"In this experiment, I want you to read aloud a series of sentences that will be displayed on this screen. There are three parts to the experiment. In the first part, you will simply read each sentence aloud. In the second part, you will read each sentence while we have you engage in a distracting task. In the third part, you will again read aloud a series of sentences, only this time, we will ask you to convey a particular emotion in your voice. I will be here with you throughout the experiment. It won't take very long. Do you have any questions?"

[Participants then read sentences for the neutral eyebrow condition.]

"Now for the second part of the experiment, we are going to make the task a little more difficult this time. While reading each sentence, I now want you to move your eyebrows, either moving them toward your nose [demonstrates] or moving them toward you hair [demonstrates].

Can you just try that for me?"

"Can you move your eyebrows toward your nose? And toward your hair? Each sentence will be accompanied by two pictures like this. In some, you will see the arrow pointing toward the nose, and in others the arrow will point toward the hair. You will need to move your eyebrows appropriately and keep them in that position while you read the displayed sentence. I will be looking at you from time-to-time to make sure you keep you eyebrows in the right position while you read the entire sentence.

Do you have any questions?"

After reading the sentences using the eyebrow condition, participants moved on to the nominal "third" part of the experiment (which was actually another experiment unrelated to the current study).

In preparing our experiment, we discovered that some people (including at least one lab colleague) were unable to move their eyebrows up or down voluntarily. For these individuals, eyebrows would move only by making specific facial expressions, like pretending to be surprised. In light of this informal observation, we screened our participants to ensure that they were able to move their eyebrows voluntarily. Of 32 recruited participants, one was excluded from the experiment because she was unable to willfully move her eyebrows.

Each participant read 12 sentences in three conditions: (1) eyebrows raised, (2) eyebrows lowered, and (3) neutral eyebrow placement. As noted, the experimenter remained in the testing room throughout the experiment in order to ensure that the participant complied with the eyebrow placement instruction. Rather than staring at the participant's face, the experimenter maintained a discrete manner, sitting beside the participant, and glancing from time-to-time at the participant to ensure that the eyebrows were properly positioned. In the vast majority of cases, participants followed the instructions regarding eyebrow placement. However, in two cases, participants tended to want to return their eyebrows to a neutral position, and so the experimenter needed to remind them to keep their eyebrows in the appropriate place.

In selecting the sentences, we aimed to satisfy four criteria. First, texts should be long enough to collect sufficient data to allow a reliable measure of average pitch height. Specifically, texts should provide at least 1 s of voiced sound. Second, the texts should be emotionally flat or neutral, so the text does not encourage participants to employ a happy, angry, or other emotionally charged prosody. Third, the texts should be simple enough that there is a low probability of participants stumbling when speaking them. Finally, we found that some texts can lead to laughter if they are regarded as trivial, tautological, or "stupid." We assembled a set of around 30 sample sentences, and pre-tested them with independent participants for ease of reading aloud. We also solicited comments from these participants in order to determine whether the sentences were deemed trivial or laughter-inducing. The 12 sample texts used are given in the Appendix. Despite our intended selection criterion to avoid emotionally charged sentences, one sentence ("Abe Lincoln was killed in 1865.") somehow was included. We discuss this sentence further in Sec. VI.

Since each sentence was read aloud three times, we were concerned with possible rehearsal effects. In pre-experiment tests, we found that readers would sometimes stumble or hesitate when reading a text, especially when reading the text for the first time. Since the critical conditions are the high and low eyebrow conditions, we recorded the neutral condition first. Each participant began by recording all 12 sentences in a unique random order with neutral eyebrow placement. The remaining trials were also randomized for each participant. The random orders were constrained so that the same text did not appear on consecutive trials. For example, a sentence with high eyebrow placement was not followed by the same sentence with low eyebrow placement. Participants were recorded using a small inconspicuous microphone placed roughly 12 in. away from the mouth.

VI. RESULTS

Prior to examining any data we created an *a priori* analysis strategy. In the first instance, we were concerned about

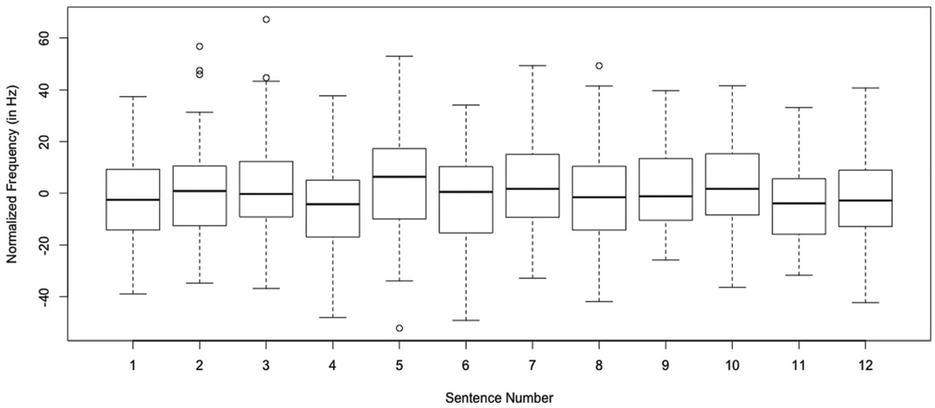

FIG. 1. Average pitch height for the 12 sentences used in the study. The vertical axis shows the frequency offset from the mean for all spoken sentences (in hertz). Data show the normalized means corrected for the differences in gender. No significant differences were observed between sentences.

the possible impact of the potentially emotionally-charged "Abe Lincoln" sentence. Despite the counterbalanced design, it is possible that an emotionally charged sentence might introduce some unanticipated confound.

Moreover, having read an emotionally charged sentence, it is possible that subsequent sentences might also be influenced after having read this sentence. Accordingly, we formulated the following *a priori* strategy. We would first analyze the average $F0$ for all 12 sentences. If the average $F0$ for sentence 3 ("Abe Lincoln") lies higher or lower than the remaining 11 sentences, we would discard all of the data for this sentence, as well as the data for the ensuing 3 sentences read by the participant. As it turned out, the $F0$ for sentence 3 was not unusually high or low compared with the other sentences, so this precautionary exclusion criterion proved unnecessary.

Post-experiment interviews were conducted with each participant. Among other purposes, the main goal of the interview was to alert the experimenters to possible unanticipated confounds, including demand characteristics. After several open-ended questions, each participant was asked to speculate about the purpose of the experiment. The majority of participants showed no evidence of having deciphered the purpose of the experiment, or of conjuring some other experimental scenario that might similarly have confounded the results. However, two participants were deemed to have accurately deciphered the purpose. One participant said, "It probably has to do with pitch changes based on where your eyebrows are. I figured this out about two examples in. I noticed the change in my voice." A second person said, "I noticed the pitch as I moved my eyebrows, I fought the tendency to change the pitch of my voice." Accordingly, the data for these two participants was excluded (*a priori*) from the data analysis.

The average $F0$ was measured for each of the 1008 recorded utterances using the Praat speech analysis software (Boersma and Weenink, 2012). Figure 1 provides box-plots for each of the 12 sentences. As can be seen, the mean $F0$ for sentence 3 does not differ from the other sentences.

The general results are summarized in Fig. 2 which plots the mean $F0$ frequency for the low, neutral, and high eyebrow conditions. The main hypothesis predicts that the pitch for low eyebrow placement will be lower than for the high eyebrow placement.

A one-way within-subject repeated measures analysis of variance (ANOVA) was conducted to compare the effect of eyebrow placement on pitch height (see Table I). After removing the effects of subject and sentence, there remains a statistically significant effect of eyebrow placement in the direction predicted by the hypothesis $(2, 967) = 15.05$, $p < 0.0001$).

Post hoc comparisons using the Tukey HSD test indicated that the main effect was significant only between Low-High and Medium-High eyebrow placement (4.91 and 3.71 Hz, respectively, with confidence bands of $\pm$2.19 Hz; $p < 0.0001$). Notice, however, that the effect size is very small (about 4 Hz) (see Table II).

Taken together, these results suggest that eyebrow placement does have an effect on vocal pitch height. Specifically, raised eyebrows cause an increase in vocal

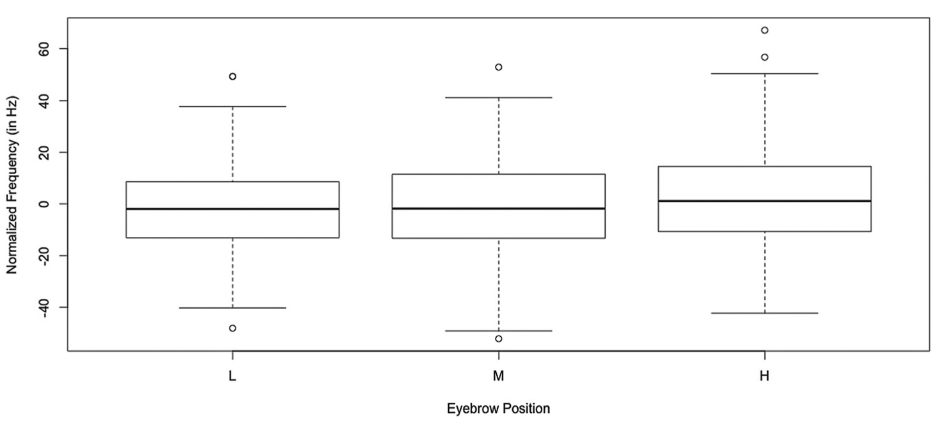

FIG. 2. The effect of eyebrow position on vocal pitch height ($F0$). Data represent the average pitch for 1008 spoken sentences. The vertical axis shows the frequency offset from the mean for all spoken sentences (in hertz). L = low eyebrow position, M = neutral eyebrow position, and H = high eyebrow position. Although the effect size is quite small, there is a statistically significant difference in mean pitch height for the high eyebrow condition compared with the neutral and low eyebrow conditions.

TABLE I. One-way within-subject ANOVA results.

	Degrees of freedom	Sum of squares	Mean Sq	F value	Pr(>F)\fR
Subject	27	1 911 130	70 783	482.373	$p < 2 \times 10^{-16}$
Sentence	11	12 068	1097	7.476	$p < 2.02 \times 10^{-12}$
Eyebrow	2	4416	2208	15.046	$p < 3.68 \times 10^{-7}$
Residuals	967	141 896	147		

pitch height. Lowering the eyebrows, however, does not cause a significant decrease in the frequency of the voice. In general, the results are consistent with the experimental hypothesis although the effect size is very small.

VII. DISCUSSION

Two results bear further discussion. First, significant results were found only in the high eyebrow condition (when contrasted with the neutral and low conditions). It is possible that this asymmetry is related to a similar phenomenon in speech prosody. Studies (reviewed in Ladd, 2008) have shown that when a speaker expands his/her pitch range, most of the change occurs in the high region: Low pitches are lowered, but not as much as high pitches are raised (see, e.g., Shriberg *et al.*, 1996).

A second notable result is the small effect size. Recall that we eliminated the data from two participants who had correctly inferred the purpose of the experiment. For example, one of the participants deduced the purpose of the experiment by noticing the change in her own voice. Excluding such data would almost certainly reduce the effect size. The very short length of the stimuli might also have contributed to a smaller effect size. Nevertheless, the small effect size in this study contrasts with the much larger effect found in Huron *et al.* (2009). At face value, moving the pitch of the voice has a greater causal impact on eyebrow placement, than vice versa. Nevertheless, the fact that there is some evidence for a two-way causality implies that a central motor process may be involved. It is possible that there exists a central "signal generator" that controls both eyebrow height and vocal pitch height.

As noted in Sec. I, any association between eyebrow height and pitch height might simply be a consequence of Hebb's Rule. However, this begs an important question. In order to "wire together," there must first be a tendency to "fire together." If we assume that eyebrow movement and vocal pitch begin as independent behaviors, we have to ask why they would tend to exhibit correlated movements that would ultimately lead them to be "wired together." That is, Hebb's Rule still leaves unanswered the prior association of eyebrow movement and vocal pitch.

Whatever the motor arrangement, the link between visual and acoustical features is consistent with the ethological notion of a signal. That is, signals are more conspicuous when they are multimodal.

VIII. TWO SIGNALS

It is interesting to note several parallels between the eyebrow/$F0$ relationship and Ohala's theory of the smile (Ohala, 2009). Phoneticians commonly characterize speech in terms of source-filter theory (e.g., Fant, 1960) in which the activity of the vocal folds ("source") is distinguished from the activity of the vocal tract ("filter"). The oscillating frequencies of the source and filter are under independent motor control. For example, high vowels (such as [i]) are distinguished from low vowels (such as [u]) predominantly by the positions of the chin and tongue producing a smaller or larger resonant cavity. Both vowels can be produced with either a low $F0$ or high $F0$, so there are four cardinal source/filter combinations: Low $F0$/low vowel, high $F0$/high vowel, low $F0$/high vowel, and high $F0$/low vowel.

At face value, there appear to be two signaling systems: One linked to the vocal source and the other linked to the vocal filter. Both putative signals combine characteristic acoustical features with characteristic visual features. One signal combines filter + mouth-configuration (i.e., smiling). The second signal combines source + eyebrows. Notice that both systems appear to derive from sound-size symbolism. That is, the acoustical components of both signals rely on conveying acoustical information suggesting "small" or "large" size.

Regarding the position of the eyebrows, Ohala (2009) has offered a speculative theory as to why high and low eyebrows might be interpreted as appeasing and threatening, respectively. Specifically, he suggested that raised eyebrows encourage retraction of the eyelids and draw attention to the eyes—effectively increasing the apparent eye-size-to-head-size ratio. Conversely, the lowered eyebrows reduce the apparent size of the eyes. Large eyes are a classic feature of infant faces—faces that are known to temper or subvert aggression and solicit compassion and nurturing (Lorenz, 1943; Glocker *et al.*, 2009). Although Ohala's conjecture

TABLE II. A post-hoc Tukey's HSD illustrating effects between eyebrow placement.

Eyebrow placements	Diff	Lower confidence band	Upper confidence band	Adjusted p-value
Low-High	−4.917909	−7.1116728	−2.724146	0.0000005
Medium-High	−3.713080	−5.9068437	−1.519317	0.0002249
Medium-Low	1.204829	−0.9889346	3.398593	0.4016135

regarding the eyebrows remains speculative, it is consistent with the ethological principle of redundancy: If signals are intended to be communicative, then the signaler should "pull out all of the stops" and employ multiple modalities to ensure successful communication.

It is possible that the two conjectured signal systems are synonymous. Both appear to have similar "meanings" pertaining to aggressive or affiliative behaviors. However, more careful descriptive work is needed in order to disambiguate the functions of these two purported signal systems. In describing the etiology of these affective displays, researchers can again benefit by understanding ethological principles. Traditionally, affective displays have been regarded as communicating or echoing the emotional state of the one signaling (e.g., Ekman and Rosenberg, 2005). However, ethological theory suggests that this view is misguided. Ethologists cogently argue that the purpose of a signal is to change the behavior of the observer, rather than to express the feeling state of the signaler (e.g., Smith and Harper, 2003). Signaling systems would not evolve if they did not ultimately enhance the inclusive fitness of the signaler. So the key to understanding the role of putative signals, such as the smile, the pout, or the raised eyebrows/pitch, lies in examining the ensuing behavioral changes in the observer, and how these behavioral changes might benefit the signaler.

Accordingly, the way to describe signals in functional terms is to employ language of the following form: Signal x tends to evoke behavior y in an observer that has benefit z for the signaler. Although this approach is widespread in ethological circles, its application to human signaling awaits.

APPENDIX: SPOKEN TEXTS

(1) Miles Davis was an American jazz trumpeter.
(2) Ohio joined the union in 1803.
(3) Abe Lincoln was killed in 1865.
(4) Most cameras don't use film anymore.
(5) The car wouldn't start this morning.
(6) Air travel is quite expensive in America.
(7) Blood pressure rises when you get stressed.
(8) My brother is ten years old.
(9) I enjoy reading a good book.
(10) Sometimes I don't get enough sleep.
(11) John Adams was the second American president.
(12) I go grocery shopping every weekend.

[1]It should be noted that brighter and darker timbre is not a consequence of VTL alone. Resonances also depend on the positions of vocal-tract constrictions with respect to the nodes and antinodes in the standing waves that underlie the resonant frequencies (see Chiba and Kajiyama, 1941/1958; Fant, 1960).

Boersma, P., and Weenink, D. (**2012**). *Praat: Doing Phonetics by Computer* (Computer program), Version 5.3.12, retrieved 4/5/2012 from http://www.praat.org/.

Bolinger, D. L. (**1964**). "Intonation across languages," in *Universals of Human Language, Vol. 2: Phonology*, edited by J. H. Greenberg, C. A. Ferguson, and E. A. Moravcsik (Stanford University Press, Stanford, CA), pp. 471–524.

Bonfiglioli, L., Caterina, R., Incasa, I., and Baroni, M. (**2006**). "Facial expression and piano performance," in *Proceedings of the 9th International Conference on Music Perception and Cognition*, pp. 1355–1360.

Chiba, T., and Kajiyama, M. (**1941/1958**). *The Vowel, its Nature and Structure* (Tokyo-Kaisekan Publishing Company, Ltd., Tokyo), 236 p.

Ekman, P., and Friesen, W. V. (**2003**). *Unmasking the Face: A Guide to Recognizing Emotions from Facial Expressions* (Malor Books, Los Altos, CA), 212 p.

Ekman, P., and Rosenberg, E. (Eds.) (**2005**). *What the Face Reveals*, 2nd ed. (Oxford University Press, New York).

Fant, G. (**1960**). *Acoustic Theory of Speech Production* (Mouton, The Hague, Netherlands), 328 p.

Glocker, M. L., Langleben, D. D., Ruparel, K., Loughead, J. W., Gur, R. C., and Sachser, N. (**2009**). "Baby schema in infant faces induces cuteness perception and motivation for caretaking in adults," Ethology **115**(3), 257–263.

Hasson, O. (**1994**). "Cheating signals," J. Theor. Biol. **167**, 223–238.

Huron, D., Dahl, S., and Johnson, R. (**2009**). "Facial expression and vocal pitch height: Evidence of an intermodal association," Empirical Musicology Rev. **4**(3), 93–100.

Johnstone, R. A. (**1997**). "The evolution of animal signals," in *Behavioural Ecology*, edited by J. R. Krebs and N. B. Davies (Oxford University Press, Oxford), pp. 155–178.

Ladd, D. R. (**2008**). *Intonational Phonology*, 2nd ed. (Cambridge University Press, Cambridge), 370 p.

Lorenz, K. (**1939**). "Vergleichende verhaltensforschung" ("Comparative behavioral research"), Zool. Anz. **12**, 69–102.

Lorenz K. (**1943**). "Die angeborenen Formen moeglicher Erfahrung" ("The innate forms of possible experience"), Z. Tierpsychol **5**, 235–409.

Morton, E. S. (**1977**). "On the occurrence and significance of motivation-structural rules in some bird and mammal sounds," Am. Nat. **111**, 855–869.

Morton, E. S. (**1994**). "Sound symbolism and its role in non-human vertebrate communication," in *Sound Symbolism*, edited by L. Hinton, J. Nichols, and J. Ohala (Cambridge University Press, Cambridge), pp. 348–365.

Ohala, J. (**1980**). "The acoustic origin of the smile," J. Acoust. Soc. Am. **68**, S33 (1980).

Ohala, J. (**1982**). "The voice of dominance," J. Acoust. Soc. Am. **72**, S66 (1982).

Ohala, J. (**1984**). "An ethological perspective on common cross-language utilization of F0 in voice," Phonetica **41**, 1–16.

Ohala, J. (**1994**). "The frequency codes underlies the sound symbolic use of voice pitch," in *Sound Symbolism*, edited by L. Hinton, J. Nichols, and J. J. Ohala (Cambridge University Press, Cambridge, UK), pp. 325–347.

Ohala, J. (**2009**). "The ethological basis of certain signals of affect and emotion," in *The Role of Prosody in Affective Speech*, edited by S. Hancil (Peter Lang, Bern), pp. 17–30.

Shriberg, E., Ladd, D. R., Terken, J., and Stolcke, A. (**1996**). "Modeling pitch range variation within and across speakers: Predicting F0 targets when 'speaking up,'" *Proceedings of the International Conference on Speech and Language Processing*, Vol. 4, pp. 650–653. Online version at www-speech.sri.com/cgi-bin/run-distill?papers/icslp96-scaling.ps.gz (Last viewed 5/9/2012), replaces version printed in proceedings.

Smith, J. M., and Harper, D. (**2003**). *Animal Signals* (Oxford University Press, Oxford), 176 p.

Tartter, V. C. (**1980**). "Happy talk: Perceptual and acoustic effects of smiling on speech," Percept. Psychophys. **27**(1), 24–27.

Wiley, R. H. (**1983**). "The evolution of communication: information and manipulation," in *Communication*, edited by T. R. Halliday and P. J. B. Slater (Blackwell, Oxford), pp. 82–113.