Phonetica *40:* 1–18 (1983)

Cross-Language Use of Pitch: An Ethological View

JOHN J. OHALA

Phonology Laboratory, Department of Linguistics,
University of California, Berkeley, Calif., USA

Abstract. Certain signaling functions of the pitch of voice are remarkably similar across languages and cultures: (1) high or rising pitch to mark questions, low or falling pitch to mark nonquestions; (2) high pitch to signal politeness, low pitch to signal assertiveness; (3) in 'sound symbolic' vocabulary, high tone used with words connoting smallness or diminutive, low tone with words connoting largeness. These patterns can be explained by the assumption that human vocal communication exploits the 'frequency code', a cross-species association of high pitch vocalizations with smallness (of the vocalizer), lack of threat, and of low pitch vocalizations with the vocalizer's largeness and threatening intent.

Introduction

There are probably few sound-meaning correlations that are more strikingly similar across languages and cultures than the association between certain pitch (fundamental frequency) contours and the linguistic categories of STATEMENT and QUESTION [HERMANN, 1942; BOLINGER, 1964, 1978; ULTAN, 1969; CRUTTENDEN, 1981]. When the same utterance can be produced as a question or statement using only intonation, i.e., without any lexical or syntactic markers, it is almost invariably the case that high or rising pitch signals the former whereas low or falling pitch the latter. In this paper I briefly sketch a new hypothesis which explains this widespread sound pattern.

The existence of a systematic cross-language correlation between pitch contours and sentence type is, on the face of it, rather surprising. In general, sound-meaning correlations are arbitrary and must be so if a potentially unbounded set of meanings is to be symbolized by strings of variable length which constitute permutations of basic sounds from a finite sound inventory. Arbitrary sound-meaning correlations

0031–8388/83/0401–0001 $ 2.75/0

are, moreover, the norm, e.g., consider the word for *cup* in various languages: English /kʌp/, Spanish /taza/, Hindi /pjala/ – and these are from languages that are genetically related! We should examine carefully and attempt to account for any violations of the tendency towards 'arbitrariness of the sign' because they may throw into relief the nature and origin of the symbolizing process that makes human language possible.

The Physiology of Question and Statements

Logically, it is reasonable to look for the origin of cross-language sound patterns in domains that are common to all speakers and all cultures at any time in history, namely, the physical world and the properties of the speech production and perception mechanisms. Accordingly, Lieberman [1967], pursuing some suggestions made by Bolinger [1964], hypothesized that the different pitch patterns for questions (strictly speaking, yes-no questions) and statements were due to the former being a syntactically 'marked' sentence form and therefore requiring phonetic expression by means of an equally 'marked' physiological mechanism for the regulation of pitch. (Justification for regarding questions as syntactically marked, vis-à-vis statements, is the fact that when intonation is not the exclusive indicator of interrogation, an extra word or particle or some kind of syntactic rearrangement of the statement form is used. Often, though, statements require no such overt marking.) Statements, on the other hand, being unmarked, would be produced with an appropriately 'unmarked' pitch-regulating mechanism. The unmarked mechanism, according to Lieberman [1967], involves no variation in the activity of those laryngeal muscles which regulate pitch so that variations in pitch would be determined solely by subglottal air pressure (P_s). P_s falls at the end of an expiration (which is what speech is superimposed on) due to the succeeding inspiratory phase of respiration and so pitch also falls. The marked physiological mechanism used to express the syntactically marked questions, however, involves increased activity of those laryngeal muscles which raise pitch. Questions, therefore, would have high or rising pitch since the laryngeally caused pitch variations would override those caused by P_s.

There are problems, both logical and empirical, with Lieberman's hypothesis. The logical basis of his claim seems to be that in (1).

(1) Marked/unmarked syntactic structures must be expressed by phonetically marked/unmarked mechanisms, respectively.

One may ask, though, if (1) is true, why it should apply only to questions? Sentences which are negative, passive, causative, imperative, etc., are also syntactically marked. Why is there not a cross-language tendency to express these in a phonetically marked way? Logically, should (1) not apply to morphological entities, too? Thus, plurals of nouns, past or future tense of verbs or any forms which are derived or inflected from other (base) forms should be expressed with phonetically marked mechanisms. This, of course, is not the case.[1] Furthermore, even if one accepted (1) as applying only in the case of questions versus statements, why should it be the case that the phonetic means of marking an utterance as a question is an increase in the activity of the pitch-raising muscles? Why not an increase in the velum-lowering (or raising) muscles or in those which constrict the pharynx, round the lips, retroflex the tongue, etc.? And if the marked mechanism must be located in the larynx, why should it not involve muscles which increase intensity, modify voice quality, etc.? There is no answer to these questions in LIEBERMAN's account but a truly explanatory hypothesis for universals of intonation *should* answer them.

The empirical difficulties which LIEBERMAN's hypothesis faces center around the issue of whether the pitch-regulating laryngeal muscles really are inactive during the production of statements and whether P_s variations could be responsible for the pitch variations observed in non-questions. LIEBERMAN sampled only pitch and P_s in his studies; the level of activity of the laryngeal muscles was guessed at.[2] If the laryngeal muscles were inactive, his data, if properly analyzed, would indicate that P_s influenced pitch at rates of from 28 to 33 Hz/cm H_2O. These values are an order of magnitude higher than estimates obtained under controlled conditions, i.e., where some efforts were made to eliminate the participation of the laryngeal muscles during the calibration [see review in OHALA, 1978]. This would lead one to suspect that the

[1] This needs to be qualified. Reduplication, which, in a sense, is a 'marked' feature, *is* often utilized to indicate 'marked' categories such as plurals of nouns, imperfectives of verbs, etc. [SAPIR, 1921, pp. 79ff.]. However, what I am referring to here is the use of specific phonetic features, articulatory or acoustic, for marked morphological or semantic categories.

[2] It is quite possible to deduce laryngeal activity from these parameters. This was done with some success, for example, by LADEFOGED [1963]. His results, however, did not point to the laryngeal muscles being inactive during the utterance of statements.

laryngeal muscles *were* active in statements as well as questions and that P_s could not be the only cause of the observed pitch variations. These suspicions were verified by physiological studies involving several languages which showed that pitch regulation in any utterance is vested primarily in the larynx [see reviews, in OHALA, 1977, 1978, 1982a; OHALA et al., 1980]. The search for an explanation for universal tendencies in intonation, then, had to continue.

Pitch and Expressive Vocalizations

An important conceptual advance in this search was the realization that the association of pitch contours with sentence types may not be an exclusively linguistic phenomenon. Rather, it may properly belong to the realm of emotional or attitudinal expression and be only partially – perhaps loosely – manifested in linguistic signals [BOLINGER, 1964; CRUTTENDEN, 1981]. Indeed, it has long been recognized that the correlation between pitch contours and sentence type breaks down when emotional, expressive, or other 'paralinguistic' factors intervene. Sentences which are pragmatically statements may be uttered with so-called question intonation, as in (2), where the pitch of voice is indicated by the curves above the text [cf. CHING, 1982].

(2) Dr. Smith. I'd like your permission to audit your course.

Similarly, yes-no questions may be uttered with a low or falling pitch as in (3), which might be a sentence a mother would address to an unruly child.

(3) Are you going to do as you are told?

The former deviation is said to express politeness, deference, or submission; the latter, aggressiveness, assertiveness, or threat. BOLINGER [1978] reviews several cases from diverse languages and cultures where pitch of voice is used with social meanings identical or similar to these, e.g., in Kunimaipa a rising pitch also signals politeness. (I will return later to the evidence for this use of pitch to convey attitude.)

Pitch in Size-Sound Symbolism

Another cross-linguistic use of pitch that might profitably be considered along with those just discussed is the patterning of tone (in tone languages) in sound symbolism.[3] As in the case of intonation, considered above, sound symbolism involves the systematic, i.e., non arbitrary association of sound and meaning in a small portion of a language's lexicon. This vocabulary includes imitative words such as Hindi [b̤õb̤õ] ([b̤] is the murmured stop) 'dog's bark' and French [kokoriko] 'rooster's crow' and such language-specific sound/meaning associations as that between the English consonant cluster [gl-] and the notion of '(brief) flash of reflected light; fleeting visual image' as manifested in such words as *glimpse, glance, glare, gleam, glimmer, glitter, glint, glisten, glow, and gloss*. Of more interest, I believe, are cases of cross-language similarities in sound/meaning correlations where the meaning does not seem to be one that is capable of imitation by speech sounds. One such case is size-sound symbolism (and related semantic dimensions such as distance, diminutive-augmentative). Several cross-language studies, some of them of a statistical or experimental nature, demonstrate the widespread tendency to use [acute] segments for words connected with the meaning SMALL, as opposed to [grave] segments with the meaning LARGE [Sapir, 1929; Jespersen, 1933; Jakobson and Waugh, 1979; Ohala, 1982b]. Some examples are given in (4).

(4) Spanish [tʃiko] 'small'	[goɾdo] 'fat'
French [pətit] 'small'	[gʁã] 'large'
Greek [mikros] 'small'	[makros] 'large'

Segments which are [acute] are those with energy predominantly in the high frequencies whereas [grave] sounds have most of their energy in the low frequencies. Although extensive cross-language surveys of how *tone* is used in size-sound symbolism have not yet been done, there is some evidence that it patterns in a way similar to that of segments, i.e., high tone being associated with SMALL and low tone with LARGE.

Westermann [1927] and Welmers [1973] provide examples of the sort in (5) from West African languages.

[3] Bolinger [1978] and Liberman [1978, p. 92] also remarked on the similarities in the use of pitch in intonation to the use of segmental phonetic features in sound symbolism.

(5)	Twi	[kàkrá] 'small'	[kàkrà] 'large'
	Ewe	[kítsíkítsí] 'small'	[g͡bàg͡bàg͡bà] 'large'
	Yoruba	[bírí] 'be small'	[bìrì] 'be large'
		[g͡bóró] 'be narrow'	[g͡bòrò] 'be wide'
		[k͡pénk͡pé] 'small'	[g͡bẹ̀ng͡bẹ̀][4] 'large'

In addition, certain Cantonese words can have their citation tone modified such that their end point is elevated to the highest tonal level when the connotation 'diminutive' or 'that familiar thing' is added [CHAO, 1947; WHITAKER, 1955/56]. Examples are given in (6).

(6)	[to 21] 'terrace, stage'	[to 215] 'table'
	[gam 33 da:i 22] 'so big'	[gam 33 da:i 335 'only so big' cf. also [gam 33 da: i 55] 'so small')
	[da:i 22 mu:n 21] 'main door'	[wa:n 21 mu:n 215] 'side door

A similar phenomenon, called 'induced creaky tone', is found in colloquial Burmese [OKELL, 1969].

An apparent reversal of the above pattern, however, is the use of high tone (instead of the expected low tone) in Bushman to convey the notion that a named object is distant or large [BLEEK, 1929]. Clearly, more research is needed on this topic but I will assume for the present that the use of pitch (tone) in size-sound symbolic words parallels the use of segments in such vocabulary, namely high frequency for SMALL words and low frequency for LARGE words.

Pitch in Nonhuman Vocalizations

Just as it was an important conceptual advance to realize that the association of pitch contours with certain sentence types was best treated as a manifestation of a much larger phenomenon, the use of pitch to signal attitude or emotional state, so I believe that our understanding of this area will be even further advanced by recognizing that

[4] The use of [± voice] in these examples is consistent with the pattern of high versus low-frequency sounds being associated with SMALL versus LARGE, respectively, since voiceless obstruents are characterized by higher frequencies than voiced obstruents. See also NICHOLS [1971].

these patterns of use of pitch are not limited to human vocalizations. They are, in fact, equally widespread (perhaps universal) in nonhuman vocalizations, too.

Morton [1977] conducted a survey of the literature on the vocalizations of 28 avian and 28 mammalian species and discovered remarkable cross-species similarities in the acoustic form of those sounds used in face-to-face competitive encounters. (This excludes mother-infant, warning, food-finding, and other vocalizations used in noncompetitive situations.) The sounds made by a confident aggressor or dominant individual are low-pitched and often harsh, whereas those of a submissive or subordinate individual are high-pitched and tonelike. The dog's threatening growl and submissive whine are familiar examples of this pattern which I will henceforth refer to as the 'frequency code.' The remarkable thing is that these polar extremes of the frequency and harshness scale are used in the same way by species as different as the chickadee and the rhinoceros and, as recent evidence suggests, frogs [Ryan, 1980]. Morton [1977] provided the following explanation for this association of the acoustic properties of the vocalization with the intent or attitude of the vocalizer:

In most cases a fight will be won by the larger of the two combatants. Since this is recognized[5] by both parties, an actual fight, which might lead to injury to both – even the victor –, can be avoided if the antagonists can inform each other of their relative sizes. Thus, the one that is smaller can yield or submit to the larger and in this way the dispute can be settled without bloodshed. For this reason animals have evolved a considerable variety of signals, mostly visual, which allow them to convey an impression of their body size or, in the case of bluffs, apparent size, e.g., threatening dogs erect the hair on their backs, raise their ears and tails, cats arch their backs, birds extend their wings and fan out their tail feathers. The pitch of voice can also indirectly convey an impression of the size of the vocalizer since there is a correlation (an inverse one) between the rate of vibration of the vocal cords (or the syringeal membranes) and overall body mass (because body mass predicts vocal cord or syringeal mass which in turn determines

[5] Here and elsewhere in this paper I take the usual ethological 'license' in using terms such as 'recognize', 'inform', and others which might seem to imply intention or human-like cognition on the part of the behaver. These are 'shorthand' expressions for the more correct formulation: A given behavior is established because those individuals which were genetically inclined to produce it had a survival and reproductive advantage over those who were not.

modal pitch). Moreover, the more massive and thick the vibrating mass, the more likely it is that secondary modes of vibrations will be set up in it and in that way give rise to irregular vibratory patterns and thus harsh voice quality. The individual that intends to prevail in the contest – initially both competitors may – will try to convey his largeness (even if it is a bluff)[6] by producing the lowest-pitched and harshest vocalization he is capable of. The individual that wants to capitulate will attempt to appear small and nonthreatening and will therefore emit as high-pitched and tonelike a cry as possible. Conceivably, it may be that in some cases such high-pitched cries constitute infant mimicry [EWER, 1968; pp. 211, 215, 232 f.; TEMBROCK, 1968]. If so, this would be an excellent strategy to ward off aggression since in most species there has evolved – for obvious reasons – a very strong taboo against harming infants.

The Explanatory Value of the Frequency Code

If we assume the frequency code is used by humans, then it helps to explain the cross-language and cross-cultural uses of pitch discussed above: its systematic correlation with sentence types, speaker attitude, and with sound symbolic words conveying notions of size.[7]

Consider, first, the pitch and sentence type connection. 'Size' is the immediate meaning projected by the pitch of voice but the ultimate meanings, by association, are undoubtedly much wider. Low pitch probably means not only 'large' but also aggressive, assertive, self-confident, dominant, self-sufficient, etc. The meaning of high pitch in addition to 'small' is nonthreatening, submissive, subordinate, in need of the receiver's cooperation and good will, etc. A person asking a

[6] Bluffs present something of a problem for evolutionary accounts of behavior: If a bluff, i.e., deceit, can evolve, than so can a way of penetrating that bluff. Then why are they so stable? It is possible that bluffs take time to penetrate and that their effectiveness is maximum only at the very initial stages of the encounter. Nevertheless this may be enough to make them beneficial to the bluffer.

[7] As alluded to above, I believe the frequency code exploits not only the pitch or fundamental frequency of vocalizations, but also their resonances, since resonant frequencies can also indirectly convey the size of the vocalizer. I have outlined elsewhere how this extension of the frequency code could explain universal patterns of size-sound symbolism using vowels and consonants [OHALA, 1982b] and could explain as well the origin of certain facial expressions such as the smile, which involve distinctive shaping of the mouth in a way that would alter the resonances of an accompanying vocalization [OHALA, 1980, 1982c].

question is implicitly dependent on others – for the information requested. To signal this using the frequency code, a high pitch would be required. A person making a statement, however, is implicitly declaring his certainty or self-sufficiency – from an information standpoint – and if conveying this by the frequency code would use low pitch. If so, we have in this way explained why these different pitches are associated with questions versus nonquestions and not with sentences which are passive, causative, etc. and why it is high pitch associated with questions and not variations in nasalization, pharyngealization, labialization, etc. The frequency code, then, although originally evolved for nonlinguistic communication, is overlaid on articulate speech such that the intention or attitude of the speaker may be further clarified. In much the same way, in face-to-face conversation, facial expressions and other forms of 'body language' can further modify, clarify, elaborate, or even contradict the semantics of the strictly linguistic message.

As for the apparent widespread use of pitch to convey attitude, the applicability of the frequency code is obvious. Expression of politeness, deference, being subordinate, on the one hand, and assertiveness, dominance, on the other, would, according to the frequency code be transmitted by high and low pitch, respectively.

Before leaving this point, though, it is necessary to discuss in more detail the evidence that humans do use pitch in this way. Common experience provides some relevant evidence. In so-called 'assertiveness training' classes, participants are told explicitly to use as low a pitch of voice as they are comfortable with in order to enhance their assertiveness or image of self-confidence. Radio and television broadcast stations tend to hire announcers whose voice pitch is lower (in comparison to the general population) so that they may sound authoritative and confident. Actors and actresses are generally 'locked into' certain types of roles as a function of their voice pitch; e.g., it would seem ludicrous to have Macbeth, Hamlet, or Richard III played by actors with high-pitched voices. There is also evidence from experimental studies in support of this point. One of the vocal cues of 'fear' (presumably related in some fashion to appeasement or submission) is high pitch [Fairbanks and Pronovost, 1939; Williams and Stevens, 1972]. Brown et al. [1974] also found that the higher the pitch of voice, the less 'competent' the speaker was perceived to be. Apple et al. [1979] found higher pitch of voice to correlate with the speaker being perceived as 'less truthful, less emphatic, and less "potent" (smaller...) and more nervous'. On the other hand, one of the cues for 'contempt' (presumably closely related to aggression or dominance) is low pitch [FAIRBANKS and PRONOVOST, 1939]. ULDALL [1960, 1964] found some tendency, occasionally a weak one, for falling pitch to lend authority to a speaker's voice and rising pitch to signal his submissiveness.

Contrary to these tendencies, SCHERER et al. [1973] found higher pitch of voice used to express confidence in some but not all circumstances. One might also suppose that the cues for 'anger' would prove to be those predicted for aggression, i. e., low pitch, but this seems not to be the case: Several studies [reviewed by SCHERER, 1979] report that high pitch and the use of a wide pitch range were the most characteristic vocal attributes of the emotional state labeled 'anger'.

We should be cautions, however, in our use and interpretation of these data since there are numerous factors not always fully recognized which might distort them or reduce their relevance to the issues discussed in this paper. There is, first of all, some question as to whether speakers can adequately simulate these different psychological states on command, even if they are professional actors. Second, what is of interest for the present discussion are those vocalizations which are intended to have some effect on a receiver, i. e., signals which are manipulative. If there are, as is often claimed, ways of using the voice just to express the emotional state of the speaker and for which a receiver is not essential, then these are not (or are not necessarily) relevant to our concerns. Because there may be these two types of signals, I do not know how to interpret the label 'anger'. Does this refer to what might be called 'frustration' where no receiver is implied, or aggression where a receiver is required? Third, there may be problems in knowing what to measure in order to show a quantitative correlation between some vocal feature and a given emotion. For example, in a pitch fall, should one measure its high point, low point, average pitch, its slope, its range? I suspect that what matters primarily to the listener is the end point of the contour [HOMBERT, 1975, pp. 102ff.] except that the 'lowness' of this low point can be enhanced by having been reached from a higher high point. If this speculation is correct, then reports that higher peak pitch is found in voices showing anger may not tell the whole story. Needless to say more research is needed on all these points.

A study by CARLETON and OHALA [1980] avoided some of these potential problems. They obtained listeners' subjective ratings of the 'attitudes' expressed in several short speech samples which had been acoustically processed such that all segmental information had been removed leaving a buzz that retained the amplitude modulations of the original and either retained the pitch of the original or had a linearly upshifted or downshifted version of the original pitch variations. They found a highly significant inverse correlation between the average pitch of the sample (whether shifted or at the original level) and judges' ratings on the scales 'dominant-submissive' ($r = -0.83$) and 'confident-unconfident' ($r = -0.88$). These results and most of those reviewed above are consistent with the notion that humans use pitch to express social intentions and attitudes.

The frequency code would give only a partial explanation for the apparent cross-language tendency to associate high tone with words signifying SMALL (and related concepts) and low tone with LARGE (see also footnote 7). What is missing from the account is how these extremes of pitch came to be used denotatively as opposed to expressively. The speaker using a high pitch to signal questions or deference is, according to the hypothesis entertained here, using the frequency code in much the same way as a submissive whining dog, namely, to convey an impression of himself which will evoke good will and cooperation from the receivers of the signal. The speaker using high pitch to denote something small is presumably not trying to convey any impression of self but rather of something distinct from self. If we could explain this different use of pitch we would have solved a major part of the problem of how human language itself came to differ from the signaling systems employed by other species. Perhaps it is sufficient for the present to account for the sound-meaning correlation only and leave for the future this more difficult question of how the correlation is used.

Supporting Evidence: Sexual Dimorphism of Human Vocal Anatomy

As suggested above, it only makes sense to invoke the frequency code as an explanation for the way pitch is used in human communication if we assume that humans do use this code. There is, I believe, strong evidence that the frequency code is an integral part of the human vocal communication system, namely, that the human vocal apparatus is 'shaped' in such a way to permit maximal exploitation of the frequency code. Specifically, this evidence is the fact that there is sexual dimorphism in the human vocal anatomy. From the onset of puberty up to approximately 20 years of age the male and female larynx undergo quite different growth patterns such that in the adult the male's vocal cords are about 50% longer than those of the female [LUCHSINGER and ARNOLD, 1965, p. 63] and his larynx as a whole is lower in the neck, thus giving him a longer vocal tract [GOLDSTEIN, 1979].[8] Because of these anatomical differences, the male adult has a lower pitch of voice and lower vowel resonances in comparison with the female adult [PETERSON and BARNEY, 1952].

Why should these differences exist? Many phonetics and speech textbooks ignore sexual dimorphism in the vocal anatomy entirely. A few note its existence but offer no speculations on its possible cause or purpose. NEGUS [1949] is a notable exception: He entertains the possibility that the larger larynx of the male permits a greater rate of air intake into the lungs and therefore permits more vigorous physical activity. However, he also points to evidence which seems to undercut this hypothesis, namely, that the male trachea and bronchi are not larger than those of the female in the same proportion as exhibited by the laryngeal opening. So even though the 'mouth of the funnel' may be larger, if the 'neck' is not also larger, the larger mouth will serve little purpose.

Although the male/female difference in the larynx has been given little attention, the low position of the human larynx vis-à-vis those of nonhuman primates has been discussed a great deal. Since the speculations on the causes of this interspecies difference have some bearing on the intraspecies sex-determined difference, a brief review and critique of these earlier hypotheses is necessary. Some have argued that the lower human larynx is a special adaptation necessitated by humans' erect posture [DUBRUL, 1976], humans' lack of a pronounced snout [NEGUS, 1949], or that it is a special adaptation for speech [LIEBERMAN, 1972]. All of these speculations ignore two crucial facts. First, the adult female larynx is *not* remarkably low in the neck. To maintain

[8] Data on the relative position of the larynx in adult males and females are hard to find. From X-ray figures such as those presented by HOLBROOK and CARMODY [1937], one can estimate that the male's vocal cords lie opposite the middle of the sixth cervical vertebra whereas the female's lie opposite the middle of the fifth. This would result in at least a 2-cm difference in overall tract length.

the speculations mentioned above one would also have to allow that human females were not as well adapted as males for erect posture, lack of a snout, or for speech. These are absurd conclusions, of course. Second, these speculations ignore the developmental aspects of the laryngeal descent in males; it takes place at puberty and continues for another 8 years or so. As a rule, morphological changes are 'timed' to take place when they are needed. Male deer, for example, do not get their antlers until they are ready to compete for mates. Human males' erect posture, lack of a snout, and their speech are not suddenly called into play just at the onset of puberty. I therefore think we can safely reject these factors as playing any important role in the shaping of human vocal anatomy.

A clue to the purpose (i.e., the adaptive value) of the enlargement of the male's larynx and vocal tract is the fact that it occurs at the same time as other secondary sexual traits appear, e.g., facial hair. They may therefore serve similar functions. Growth of hair at the perimeter of the face in males, as with so many other primate species, increases the apparent size of the head and thus enhances the visual effectiveness of a threat display [Guthrie, 1970]. The enlargement of the vocal anatomy gives the male a lower-pitched voice and lower resonances which, according to the frequency code, also indirectly enhance his apparent size. It would seem plausible, then, that the dimorphic aspects of the vocal apparatus serve to improve the male's ability to protect the family unit, i.e., that they are adaptations to the sex-specific role of the human male (in earlier days, if not at the present stage of human society). This is consistent with current ethological speculation on the factors leading to sexual dimorphism [Ralls, 1977] and with the role played by adult males in other primate groups [DeVore, 1963]. Clearly, the differences in male and female vocal anatomy are manifestations of an innate disposition (which must be actuated by hormones, of course[9]); if these innate morphological characteristics are motivated by the frequency code, it is safe to conclude that the frequency code is an inherent part of human communication[10].

[9] Apropos of this, it is intriguing to note that Meuser and Nieschlag [1977] found a greater concentration of testosterone, the male sex hormone, in basses and baritones than in tenors.

[10] It may be worth noting in this connection that parts of the vocal anatomy that cannot directly create a marked concentration of acoustic energy in the high or low frequencies, e.g., the velum, do *not* exhibit sexual dimorphism [Subtelny, 1957].

The reasonableness of this conclusion is reinforced by the fact that many other species also exhibit sexual dimorphism in the vocal anatomy, i.e., differences in the size of the sound-producing organs that is out of proportion to any overall differences between the sexes in body mass. In many cases there is evidence that the larger vocal structures improve the animal's threat display. In elephant seals, for example, the male has a large proboscis, completely lacking in females and juveniles, that is part of its phonatory apparatus; the length of the proboscis correlates inversely with the dominant frequency of the phonation but correlates directly with the success of maintaining a harem in the face of competition from other males [BARTHOLOMEW and COLLIAS, 1962; SHIPLEY, 1981, and personal commun.]. In many waterfowl, e.g., wood ducks, mergansers, the male, but not the female, possesses a bulla, an asymmetric enlargement of the trachea just above the syrinx which probably serves to make the male's phonation different, at least louder, than that of the female [TERRES, 1980, p. 79]. In whooping cranes, the length of the trachea is about as long as the bird itself; several convolutions of the trachea are stuffed inside the hollow sternum. This serves to make their calls extremely loud – whence their name. There is evidence that the males have a longer trachea than females [ROBERTS, 1880] and that the males' calls have a lower frequency than those of females [ARCHIBALD, 1976]. One of the most extreme cases of enlargement of the vocal cavities is present in the bird of paradise *Phonygammus*, which although only about 25 cm long itself, has a trachea over 80 cm long, the 'extra' length being coiled up between the front surface of the sternum and the external skin. Females of the species have shorter tracheae than males [CLENCH, 1978]. Many primates also exhibit dimorphism in their vocal structures whether in the position of the larynx (e.g., the gorilla) or in the size of the ventricular sacs (e.g., siamang) [VALLOIS, 1955].

(To return to an earlier point, it should be noted that none of these cases of extreme enlargement of the vocal structures can be explained as constituting 'special adaptations to speech'. Furthermore, most of them cannot be accounted for as adaptations to a reduced snout and some resist the 'adaptation to erect posture' hypothesis. There is, then, little reason to invoke these explanations for the quite comparable anatomical changes in human males.)

It does not seem reasonable to suppose that the sex differences in the vocal structures of humans should exist for different reasons than comparable differences in other species. I therefore assume that they both exist for the same reason, which is, plausibly, to augment the male's defense and threat capabilities.

Conclusion

The frequency code is a cross-species sound/meaning correlation whereby vocalizations consisting of high frequencies signal the vocalizer's apparent smallness and, by extension, his nonthreatening, submissive, or subordinate attitude and by which low-frequency vocalizations signal apparent largeness and thus threat, dominance, self-confidence. The existence of sexual dimorphism in the vocal anatomy of adult humans suggests that the frequency code is an inherent part of human vocal communication (and probably has been for millions of years). The adult male's larger and lower larynx gives him the low-pitched and low-resonance voice that would enhance the vocal component of threat displays and thus permit better defense of the family unit. The frequency code explains the similarities in cross-language and cross-cultural use of the pitch of voice to mark questions versus non-questions, to signal different social attitudes (dominance, submission, assertiveness, politeness), and to refer to things small and large using sound symbolic vocabulary.

Acknowledgements

I gratefully acknowledge the support of the Harry Frank Guggenheim Foundation for the research reported here. I also thank E. Morton and J. Wheatley for advice and bibliographical tips. I assume responsibility for the ideas and the accuracy of the statements presented.

References

Apple, W.; Streeter, L.A.; Krauss, R.M.: Effects of pitch and speech rate on personal attributions. J. Pers. soc. Psychol. *37:* 715–727 (1979).

Archibald, G.W.: The unison call of cranes as a useful taxonomic tool; doct. diss. Cornell University (1976).

Bartholomew, G. A.; Collias, N. E.: The role of vocalization in the social behavior of the elephant seal. Anim. Behav. *10:* 7–14 (1962).

BLEEK, D. F.: Bushman grammar. Z. Eingeborenen-Sprachen *19:* 81–98 (1929).

BOLINGER, D.: Intonation as a universal; in LUNT Proc. 9th Int. Congr. Linguists, 1962, pp. 833–844 (Mouton, The Hague 1964).

BOLINGER, D.: Intonation across languages; in GREENBERG Universals of human language, pp. 471–524 (Stanford University Press, Stanford 1978).

BROWN, B. L.; STRONG, W.J.; RENCHER, A. C.: 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. J. acoust. Soc. Am. *55:* 313–318 (1974).

CARLETON, M.; OHALA, J.J.: The effect of pitch of voice on perceived personality traits. Annu. Meet. Kroeber Anthropol. Soc., Berkeley 1980.

CHAO, Y. R.: Cantonese primer (Harvard University Press, Cambridge 1947).

CHING, M. K. L.: The question intonation in assertions. Am. Speech *57:* 95–107 (1982).

CLENCH, M. H.: Tracheal elongation in birds-of-paradise. Condor *80:* 423–430 (1978).

CRUTTENDEN, A.: Falls and rises: meanings and universals. J. Ling. *17:* 77–91 (1981).

DEVORE, I.: Comparative ecology and behavior of monkeys and apes; in WASHBURN Classification and human evolution. Viking Fund Publ. No. 37, pp. 301–310 (Wenner-Gren Foundation, New York 1963).

DUBRUL, E. L.: Biomechanics of speech sounds; in HARNAD, STEKLIS, LANCASTER Origin and evolution of language and speech. Ann. N.Y. Acad. Sci. *280:* 631–642 (1976).

EWER, R. F.: Ethology of mammals (Plenum, New York 1968).

FAIRBANKS, G.; PRONOVOST, W.: An experimental study of the pitch characteristics of the voice during the expression of emotion. Speech Monogr. *6:* 87–104 (1939).

GOLDSTEIN, U. G.: Modelling children's vocal tracts; in WOLF, KLATT Speech communication papers, pp. 139–142 (Acoustical Society of America, New York 1979).

GUTHRIE, R. D.: Evolution of human threat display organs. Evol. Biol. *4:* 257–302 (1970).

HERMANN, E.: Probleme der Frage. Nachr. Akad. Wiss. Göttingen, Philologisch-Historische Klasse, Nr. 3/4 (1942).

HOLBROOK, R.T.; CARMODY, F.J.: X-ray studies of speech articulations. Univ. Calif. Publ. Mod. Philology, vol. 20, No. 4 (1937).

HOMBERT, J.-M.: Towards a theory of tonogenesis: an empirical, physiologically and perceptually-based account of the development of tonal contrasts in languages; doct. diss. University of California, Berkeley (1975).

JAKOBSON, R.; WAUGH, L. R.: The sound shape of language (Indiana University Press, Bloomington 1979).

JESPERSEN, O.: Symbolic value of the vowel I; in Linguistica. Selected papers in English, French and German, pp. 283–303 (Levin & Munksgaard, Copenhagen 1933).

LADEFOGED, P.: Some physiological parameters in speech. Lang. Speech *6:* 109–119 (1963).

LIBERMAN, M.: The intonational system of English (Indiana University Linguistic Club, Bloomington 1978).

LIEBERMAN, P.: Intonation, perception and language (MIT Press, Cambridge 1967).

LIEBERMAN, P.: On the evolution of human language; in RIGAULT, CHARBONNEAU Proc. 7th Int. Congr. Phonetic Sci. 1971, pp. 258–272 (Mouton, The Hague 1972).

LUCHSINGER, R.; ARNOLD, G. E.: Voice, speech, language (Wadsworth, Belmont 1965).

MEUSER, W.; NIESCHLAG, E.: Sexualhormone und Stimmlage des Mennes. Dt. med. Wschr. *102:* 261–264 (1977).

MORTON, E.W.: On the occurrence and significance of motivation – structural rules in some bird and mammal sounds. Am. Nat. *111:* 855–869 (1977).

NEGUS, V. E.: The comparative anatomy and physiology of the larynx (Hafner, New York 1949).

NICHOLS, J.: Diminutive consonant symbolism in Western North America. Language *47:* 826–848 (1971).

Ohala, J.J.: The physiology of stress; in Hyman Studies in stress and accent. S. Calif. Occ. Papers Ling., No. 4, pp. 145–168 (1977).

Ohala, J.J.: Production of tone; in Fromkin Tone: a linguistic survey, pp. 5–39 (1978).

Ohala, J.J.: The acoustic origin of the smile. J. acoust. Soc. Am. *68:* 33 (1980).

Ohala, J.J.: Physiological mechanisms underlying tone and intonation; in Fujisaki, Gårding Preprints working group on intonation. 13th Int. Congr. Linguists, Tokyo 1982, pp. 1–12 (1982a).

Ohala, J.J.: The phonological end justifies any means. Preprints of the plenary session papers. 13th Int. Congr. Linguists, Tokyo 1982b, pp. 199–208 (ICL Editorial Committee, Tokyo 1982b).

Ohala, J.J.: The frequency code and its effect on certain forms of speech and facial expressions; in House Proc. Symp. on Acoust. Phonetics Speech Modeling, paper 23/81, pp. 1–31 (Institute for Defense Analysis, Princeton 1982c).

Ohala, J.J.; Riordan, C.J.; Kawasaki, H.: Investigation of pulmonic activity in speech. Rep. Phonology Lab, Berkeley *5:* 89–95 (1980).

Okell, J.: A reference grammar of Colloquial Burmese. Part I (Oxford University Press, London 1969).

Peterson, G. E.; Barney, H. L.: Control methods used in a study of vowels. J. acoust. Soc. Am. *24:* 175–184 (1952).

Ralls, K.: Sexual dimorphism in mammals: avian models and unanswered questions. Am. Nat. *111:* 917–938 (1977).

Roberts, T. S.: The convolution of the trachea in the Sandhill and Whooping Cranes. Am. Nat. *14:* 108–114 (1880).

Ryan, M.J.: Female mate choice in a neotropical frog. Science *209:* 523–525 (1980).

Sapir, E.: Language. An introduction to the study of speech (Harcourt, Brace, New York 1921).

Sapir, E.: A study in phonetic symbolism. J. exp. Psychol. *12:* 225–239 (1929).

Scherer, K. R.: Nonlinguistic vocal indicators of emotion and psychopathology; in Izard Emotions in personality and psychopathology, pp. 495–529 (Plenum, New York 1979).

Scherer, K. R.; London, H.; Wolf, J.J.: The voice of confidence: paralinguistic cues and audience evaluation. J. Res. Personality *7:* 31–44 (1973).

Shipley, C. O.: Development of vocalization on Northern elephant seal bulls; doct. diss. University of California, Los Angeles (1981).

Subtelny, J. D.: A cephalometric study of the growth of the soft palate. Plastic reconstr. Surg. *19:* 49–62 (1957).

Tembrock, C.: Land mammals; in Sebeok Animal communication, pp. 338–404 (Indiana University Press, Bloomington 1968).

Terres, J. K.: The Audubon Society encyclopedia of North American birds (Knopf, New York 1980).

Uldall, E.: Attitudinal meanings conveyed by intonation contours. Lang. Speech *3:* 223–234 (1960).

Uldall, E.: Dimensions of meaning in intonation; in Abercrombie, Fry, MacCarthy, Scott, Trim In honour of Daniel Jones, pp. 271–279 (Longmans, London 1964).

Ultan, R.: Some general characteristics of interrogative systems. Working Papers Lang. Universals, Stanford *1:* 39–63a (1969).

Vallois, H.: Ordre des primates. Traité de zoologie. Anatomie, systématique, biologie, vol. 17, part 2, pp. 1854–2206 (Masson, Paris 1955).

Welmers, W. E.: African language structures (University of California Press, Berkeley 1973).

Westermann, D.: Laut, Ton und Sinn in westafrikanischen Sudan-Sprachen, Festschrift Meinhof, pp. 315–328 (Hamburg 1927).

WHITAKER, K. P. K.: A study on the modified tones in spoken Cantonese, Asia maj. *5:* 9–36; 184–207 (1955/56).

WILLIAMS, C. E.; STEVENS, K. N.: Emotions and speech: some acoustic correlates. J. acoust. Soc. Am. *52:* 1238–1250 (1972).

Received: October 18, 1982; accepted: October 25, 1982

Prof. JOHN J. OHALA, Phonology Laboratory, Department of Linguistics, University of California, *Berkeley*, CA 94720 (USA)