Sound Symbolism

John J. Ohala
University of California, Berkeley

Sound symbolism is the term for a hypothesized systematic relationship between sound and meaning (Hinton, Nichols, & Ohala 1994). The idea that there might be a non-arbitrary relationship between the physical aspect of a speech signal and its meaning is quite an old idea, dating back at least to the time of Plato who, in his work *Cratylus*, had Socrates debating with two pupils the issue of whether the names for things are arbitrary or whether instead they are a natural reflection of the things named. There has been considerable speculation and anecdotal evidence on this point from then up to the present time (de Brosses 1765; Jespersen 1933). I believe that phonetic studies -- combined with principles emerging from ethology, the science of comparative behavior -- can provide a novel basis for further informed speculation on this issue. Moreover, the principles applied to this topic can also elucidate several other controversial topics including the cause of universal patterns in intonation, the origin of certain facial expressions involving the mouth, the reason for sexual dimorphism of the vocal anatomy in humans (and other species), and whether the human vocal anatomy shows any adaptation special to speech.

1. Linguistic data

I begin with certain fundamental observations that have been made repeatedly in the linguistic literature. First, it is generally accepted that most of the meanings conveyed by speech signals show a purely arbitrary sound-to-meaning relation. This is one of the unassailable tenets of structural linguistics. Indeed, it could hardly be otherwise if such an essentially unlimited set of meanings is to be conveyed by a decidedly small and finite inventory of speech sounds. The large inventory of meanings, the lexicon, is created by arbitrary combinations and permutations of the members of this small set of speech sounds. But could there not be a small part of the set of meanings that have a non-arbitrary relation to the sounds expressing them? If we find similar sounds associated with similar meanings in a sufficiently diverse number of languages, especially those not genetically or areally connected, then we might with some confidence consider whether a non-arbitrary sound-meaning relation exists and then seek an explanation for it.

1.1 Cross-language patterns of intonation

In many languages it is possible to turn a statement into a question simply by varying the intonation. In these cases the question typically has a higher fundamental frequency (F0) (“pitch”) than the statement somewhere in the sentence -- at the end (as in English), on the tonic syllable, or in some cases with a greater overall upshift of F0 (Hermann 1942, Bolinger 1978). This association of higher F0 with questions and lower F0 with statements is far more common than could arise by chance.

A related use of voice F0, though not as well documented as the case with question-statement, is the use of high F0 to show politeness or deference to the person spoken to. Opposite to this is the use of low F0 to express anger, aggression, or threat. In fact, this pattern may in many cases supersede that noted above for question-statement in that utterances that are semantically and pragmatically statements have been observed to end in a rising F0 apparently in cases where the speaker can be construed as expressing lesser social standing or lesser confidence (Ching 1982).

1.2 Cross-language patterns of size sound symbolism

There is abundant literature, some of it experimental or statistical, supporting the idea that across languages there are phonetically natural classes of speech sounds that are systematically associated with expressions of size (however, see Diffloth 1994, Bauer 1996 for counter-evidence).

Some examples from a variety of languages are presented in Tables 1and 2 (for references, see Ohala 1984, 1994).

Table 1. Examples of words and morphemes expressing “small”.

Language	Word, Morpheme	Translation
Ewe	[kítsíkítsí]	‘small’
Yoruba	[bírí]	‘be small’
Yoruba	[k͡péŋk͡pé]	‘small’
Spanish	[ˈtʃiko]	‘small’
Greek	[mikros]	‘small’
English	[-i]	diminutive suffix
Spanish	[-it-]	diminutive suffix
Irish	[-in]	diminutive suffix
French	[pətit]	‘small’

Table 2. Examples of words and morphemes expressing “large”.

Language	Word, Morpheme	Translation
Ewe	[ɡ͡bàɡ͡bàɡ͡bà]	‘large’
Yoruba	[bȉrȉ]	‘be large’
Yoruba	[ɡ͡bè̞ŋɡ͡bè̞]	‘large’
Spanish	[ˈɡoɾdo]	‘fat’
Greek	[makros]	‘large’
French	[ɡʁɑ̆]	‘large’

Further data of this sort (as well as some couterexamples[1]) may be found in Hinton et al. 1994. Based on data of this sort, it has been claimed that the following sound types are predominant in the expression of “small”: high front vowels like [i ɪ y e], [-grave] consonants, voiceless consonants, high tone etc., and “large”: low back vowels like [ɑ ʌ ɔ o], [+grave] consonants, voiced consonants, and low tone. There is support for this pattern from experimental (Sapir 1929; Fischer-Jørgensen 1968) and statistical studies (Chastaing 1958, Thorndike 1945, Ultan 1978). The phonetic generalization that can be made is that the expression of size utilizes speech sounds whose characteristic acoustic frequencies vary inversely with size of the thing designated.

2. Non-human agonistic vocalizations

Some light is shed on the above patterns from a study by Morton (1977) who examined the kinds of vocalizations some 28 avian and 28 mammalian species use in agonistic displays, i.e., signals given in face-to-face competitive encounters. In all cases it was found that the confident aggressor or threatener emitted a low-pitched vocalization whereas the submissive nonthreatening individual produced a high-pitched sound. The dog’s aggressive “growl” and submissive “yelp” are familiar examples. Moreover, Morton provided a rationale for this apparently consistent cross-species pattern of sound shape and meaning or intent. One of the important determinants of success in combat is the relative size of the combatants: the larger animal (which is often the older, more mature, individual) usually wins. Accordingly, in order to avoid the adverse consequences of actual combat (where there are costs to both winner and loser), both parties put on displays, both visual and acoustic, which help to convey an impression, often a false one, of their size. Then the contest can usually be settled without actual bloodshed when one party captiulates to the other.

Erection of hair and feathers, elevation of ears, tail, wings, tail feathers, etc. are visual elements of threat displays which make the signaler subtend a larger angle in the viewer's visual field. The submissive or non-threatening animal does the opposite in order to appear as small as possible: ears, tail, wings, hair, feathers are flattened against the body or tucked out of sight. Some aspects of non-threat displays are said to contain elements of infant mimicry in order to invoke the powerful within-species inhibition against harming infants, e.g., the submissive dog's act of rolling on its back and exposing its underside. Morton points out that the acoustic component of such displays can also help to convey an impression of relative size. Normally, the rate of vibration of vocal cords or syringeal membranes is inversely related to the mass of the vibrating tissue. Accordingly an animal can "seem" large if it uses the lower F0 that would be characteristic of a larger individual with more massive vocal cords or syrinx. Likewise, an animal can seem small if it uses a high F0.

An advantage of Morton's scenario is that this behavior found in agonistic displays clearly gives the animals an advantage and thus is likely to have evolved by natural selection, i.e., to be genetic, as would be predicted by Darwinian theory -- the dominant paradigm in ethology today.

I have proposed a minor extrapolation of Morton's theory in proposing that it applies as well to some aspects of human vocalization, namely, to the use of F0 to express deference and aggression, as mentioned above, and to the use of F0 in questions vs. statements. The person asking a question can be viewed as requiring the cooperation of the person to whom the question is addressed. Therefore a supplicating intonation is appropriate. A declarative statement, on the other hand, signals the speaker's self-confidence and control of the information conveyed.

As a somewhat more substantial extrapolation of Morton's theory, I have also proposed that it is not only the F0 of the vocalization which can convey an impression of the size and thus the threat vs. non-threat intent of the vocalizer but also the frequency content of the *resonances* of the vocalization. The characteristic frequencies of the vocal tract resonances are roughly inversely correlated with the length of the vocal tract which, in turn, is correlated with the linear dimensions of the vocalizer. So acoustically high resonances should convey an impression of a small vocalizer and low resonances, of a larger vocalizer.

Does this help to explain the association of the phonetic classes of sounds noted above with words meaning "small" and "large"? On the surface it does not because the speaker uttering "small" words like the English *teeny* [ˈtʰini], does not himself wish to appear small; rather he wants to convey the smallness of the thing which that word describes. But insofar as a word such *teeny* is supposed to evoke in the listener an impression of something small, the particular sounds which the word consists of may enhance that evocation. In any case, I offer the hypothesis that cross-language patterns of size sound symbolism ultimately derive from this factor.

I have labeled the cross-language and cross-species use of acoustic frequency – whether in the sound source or the sound resonances -- to convey impressions of size (and related concepts) the "frequency code".

3. The acoustic origin of the "smile"

The frequency code can help to resolve a long-standing puzzle about the origin of the smile – which, by the way, is also manifested widely across different cultures and across different species, at least those with a movable facial cover such as canids and primates. The puzzle is this: why should a non-threatening facial display show so many teeth, which are potential weapons and thus more appropriate for a threat display? Also, in many species (I think humans included) there is another facial display which is part of threat signals which is opposite to the smile in that it involves protrusion of the lips and thus greater concealment of the teeth (I call this the 'o-face'). Even the so-called "snarl" which involves display of the canine teeth -- certainly potent weapons -- still involves a somewhat "square" mouth shape in which the mouth corners cover most of the teeth further back than these sharp teeth. Some elaborate and colorful evolutionary scenarios have been proposed to explain how a dramatic display of teeth could come to stand for an affinitive signal (see Ohala 1984, 1994 for relevant references).

I propose instead that the smile -- and its opposite, the 'o-face' -- originally served as a component of the acoustic element of these displays. In the smile the mouth corners are drawn back to effectively shorten the vocal tract and to give rise to higher resonant frequencies. This shortening is much more effective in species that have a snout, i.e., a mouth that protrudes from the plane in which the eyes are set. Most primates and certainly canids have snouts; the human species used to have something of a snout and it is a reasonable guess that these facial expressions existed then and have been retained even after homo sapiens lost its snout. The o-face, on the other hand, through mouth constriction and/or lip protrusion serves to lower the resonant frequencies.

A possible objection to this account is that humans and some primates often exhibit the smile with mouth closed or without an accompanying vocalization -- thus undercutting my claim that these mouth shapes serve an acoustic end. However, it is possible (and there are many parallels) for an activity serving a given purpose to become "ritualized" and thus performed in the absence of its original purpose or even to be performed for a completely different purpose. Today the smile may very well serve in addition as an independent visual signal. My claim is that in its origin it served an acoustic end. Although the smile and o-face may be given silently, when there *is* a non-threat or threat vocalization, these facial shapes invariably accompany them in the predicted way.

As a test of this hypothesis, Bauer (1987) found a correlation in the predicted direction between mouth opening and F0 of vocalization in chimpanzee threat and non-threat displays. There is no physiological necessity for such a correlation -- the muscles serving facial expressions do not interact with those modulating vocal cord tension -- so if it exists, it must be for the behavioral reasons adduced.

4. Sexual dimorphism of the vocal anatomy in humans

It is well known that adult male and female humans have different average dimensions of parts of their vocal anatomy. Males have a vocal tract that is about 15 to 20% longer than females (this is due to the larynx being placed lower in the neck) and they have vocal cords that are about 50% longer (and considerably more massive). These two differences account for the lower resonant frequencies of males' vowels and the nearly 1 octave difference in typical male and female F0s. But I have seen few inquiries or guesses as to *why* this is the case. It can't be due to the overall differences in linear dimensions of the males and females because this difference is usually only about 5%.

I propose that these differences exist in order to enhance the acoustic component of the adult male's threat display. Males are recognized to be the more aggressive of the two sexes and their greater production of the hormone testosterone is implicated in this. Thus males' longer vocal tract and longer vocal cords are components of the roles they play as males in the human family unit and the human social unit. Presumably these anatomical features increase a male's chances of success in garnering scarce resources -- food, territory -- for the raising of offspring. There may also be an element of sexual selection in the evolutionary development of these anatomical features since it is known that females prefer as mates those with more masculine traits such as a low-pitched voice.

In this sense, the sexually dimorphic aspects of the vocal anatomy take their place alongside similar dimorphic aspects that serve in visual displays, e.g., beards. Beards in humans serve to make the head appear larger than it really is and thus the beard owner more potent.

5. The larynx and the evolution of speech

The analysis just presented of the reason for the lowered larynx in adult males undercuts proposals that the lowered larynx is somehow an adaptation for human speech, i.e., to enable to humans to produce better vowels (Lieberman 1972). The arguments against this view are as follows:

- The larynx is low in adult males; it is *not* that low in adult females or in children, the latter of whom still manage to make intelligible vowels. To claim that the lowered larynx in males is an adaptation for speech is to claim that females are less well adapted for speech. This, of course, is nonsense; if anything, females have been shown to be better at speech and language skills than males. Females also suffer from fewer language and speech related disorders such as stuttering, dyslexia, and autism.
- The lowered larynx in males (vis-à-vis females) appears at puberty, long after the individual has started speaking. Typically, sex- or age-based dimorphic traits appear when the individual, by virtue of the role it plays, *needs* them, e.g., the camouflaging spots on the back of a fawn, or the "egg tooth" seen in some birds just before hatching out of the shell. By age two years humans have fairly well developed speech skills but their larynx is not remarkably low in the neck. The fact that this larynx lowering is accelerated in the male at puberty suggests that it is related to the male's sex-based role. No one has ever suggested that that is talking! (Related to this is that the appearance of the beard also appears at puberty in the male. I claim this is not coincidence; they both serve the same function -- one in the acoustic, the other in the visual component of threat displays.)
- Many other species besides humans show (a) a species-specific enlargement of the vocal anatomy and (b) in some cases sexual dimorphism in this enlargement. Cranes and geese are well known to have disproportionately long tracheas (which for them is part of the "vocal tract"). They are known to be able to emit much louder calls as a result. The howler monkey (well named, since its call can be heard over miles even in the dense jungle) is endowed with an enlarged and hollow hyoid bone with serves as an added resonator. Only male howler monkeys have this anatomical development. (For other examples, see Ohala 1984, 1994.) Obviously, none of these other species possess speech. I suggest that their vocal tract enhancements, including sexually-dimorphic ones, serve the same purpose as that in humans: to facilitate the vocal component of threat and non-threat displays.

6. Conclusion

I have proposed that sound symbolism, the non-arbitrary connection between certain classes of speech sounds and the meanings of the words or morphemes they are in, has a biological evolutionary origin. Briefly, there is cross-species recognition and exploitation of *the frequency code*: the association of high acoustic frequency with smallness and low acoustic frequency with largeness. The frequency code also gives us a new vantage point from which to view and explain certain facial expressions involving the mouth, sexual dimorphism of the adult human vocal tract, and the proposed link between lowered larynx and the evolution of speech.

Notes

[1] Korean is often cited as a language which has a pattern of size sound symbolism that is phonetically opposite to the one described here, e.g., /piŋkɯl/ "turning in circles", /pɛŋkɯl/ "turning in smaller circles" (Kim 1977). However, it is not clear whether such vowel ablaut patterns in Korean sound symbolism signifies "size" as opposed to "intensification". If the latter, then this pattern is not necessarily an exception to that found in most other languages.

References

Bauer, H. R. 1987. "The frequency code: oral-facial correlates of fundamental frequency," *Phonetica* 44, 173-191.

Bauer, L. 1996. "No phonetic iconicity in evaluative morphology," *Studia Linguistica* 50, 189-206.

Bolinger, D. L. 1978. "Intonation across languages," In J. H. Greenberg, C. A. Ferguson, and E. A. Moravcsik (eds.), *Universals of human language*, *Vol. 2: Phonology*. Stanford, CA: Stanford University Press. 471-524.

de Brosses, C. 1765. *Traité de la formation méchanique des langues, et de principes physiques de l'étymologie*. 2 vols. Paris: Chez Saillant, Vincent, Desaint.

Chastaing, M. 1958. "Le symbolisme des voyelles: significations des 'i'," I, II. *J. de Psychologie* 53, 403-423; 461-481.

Ching, M. K. L. 1982. "The question intonation in assertions," *American Speech* 57, 95-107.

Diffloth, G. 1994. "*i*: big, *a*: small," In L. Hinton, J. Nichols, and J. J. Ohala (eds.), *Sound symbolism*. Cambridge: Cambridge University Press. 107-114.

Fischer-Jørgensen, E. 1978. "Perceptual dimensions of vowels," *Z. f. Phonetik, Sprachwissenschaft u. Kommunikationsforschung* 21, 94-98.

Hermann, E. 1942. *Probleme der Frage*. Nachrichten v. d. Akademie der Wissenschaften in Göttingen. Philologische-Historiche Klasse, Nr. 3/4.

Hinton, L., J. Nichols, & J. J. Ohala (eds.), *Sound symbolism.* Cambridge: Cambridge University Press.

Kim, K.-O. 1977. Sound symbolism in Korean. *J. Linguistics* 13, 67-75.

Jespersen, O. 1933. "Symbolic value of the vowel *i*," In O. Jespersen, *Linguistica: Selected papers in English, French and German*. Copenhagen: Levin and Munksgaard. 283-303.

Lieberman, P. 1972. "On the evolution of human language," In A. Rigault and R. Charbonneau (eds.), *Proc. 7th Internat. Congr. of Phonetic Sciences*, *1971*. The Hague: Mouton. 258-272.

Morton, E. W. 1977. "On the occurrence and significance of motivation-structural rules in some bird and mammal sounds," *American Naturalist* 111, 855-869.

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

Ohala, J. J. 1994. "The frequency codes underlies the sound symbolic use of voice pitch," In L. Hinton, J. Nichols, & J. J. Ohala (eds.), *Sound symbolism.* Cambridge: Cambridge University Press. 325-347.

Sapir, E. 1929. "A study in phonetic symbolism," *J. of Experimental Psych*. 12, 225-239.

Thorndike, E. L. 1945. "On Orr's hypothesis concerning the front and back vowels," *British Journal of Psychiatry* 36, 10-14.

Ultan, R. 1978. "Size-sound symbolism," In J. H. Greenberg, C. A. Ferguson, and E. A. Moravcsik (eds.), *Universals of human language*, *Vol. 2: Phonology*. Stanford, CA: Stanford University Press. 527-568.

Department of Linguistics
University of California
Berkeley, CA 94720
USA
e-mail: ohala@cogsci.berkeley.edu
Fax: +1 510 649 0776