

# Essential tones of music rooted in human speech

May 24 2007

---

The use of 12 tone intervals in the music of many human cultures is rooted in the physics of how our vocal anatomy produces speech, according to researchers at the Duke University Center for Cognitive Neuroscience.

The particular notes used in music sound right to our ears because of the way our vocal apparatus makes the sounds used in all human languages, said Dale Purves, the George Barth Geller Professor for Research in Neurobiology.

It's not something one can hear directly, but when the sounds of speech are looked at with a spectrum analyzer, the relationships between the various frequencies that a speaker uses to make vowel sounds correspond neatly with the relationships between notes of the 12-tone chromatic scale of music, Purves said.

The work appeared online May 24 in the *Proceedings of the National Academy of Sciences*. (Download at <http://www.pnas.org/cgi/reprint/0703140104v1>)

Purves and co-authors Deborah Ross and Jonathan Choi tested their idea by recording native English and Mandarin Chinese speakers uttering vowel sounds in both single words and a series of short monologues. They then compared the vocal frequency ratios to the numerical ratios that define notes in music.

Human vocalization begins with the vocal cords in the larynx (the Adam's apple in the neck), which create a series of resonant power peaks in a stream of air coming up from the lungs. These power peaks are then modified in a spectacular variety of ways by the changing shape of the soft palate, tongue, lips and other parts of the vocal tract. Our vocal anatomy is rather like an organ pipe that can be pinched, stretched and widened on the fly, Purves said. English speakers generate about 50 different speech sounds this way.

Yet despite the wide variation in individual human anatomy, the speech sounds produced by different speakers and languages produce the same variety of vocal tract resonance ratios, Purves said.

The lowest two of these vocal tract resonances, called formants, account for the vowel sounds in speech. "Take away the first two formants and you can't understand what a person is saying," Purves said. The frequency of the first formant is between 200 and 1,000 cycles per second (hertz) and the second formant between 800 and 3,000 hertz.

When the Duke researchers looked at the ratios of the first two formants in speech spectra, they found that the ratios formed musical relationships. For example, the relationship of the first two formants in the English vowel /a/, as in "bod," might correspond with the musical interval between C and A on a piano keyboard.

"In about 70 percent of the speech sounds, these ratios were bang-on musical intervals," Purves said. "This predominance of musical intervals hidden in speech suggests that the chromatic scale notes in music sound right to our ears because they match the formant ratios we are exposed to all the time in speech, even though we are quite unaware of this exposure."

No music, except modern experimental pieces, uses all 12 tones. Most

music uses the 7-tone or diatonic scale to divide octaves, and much of folk music uses five tones. These preferences correspond to the most prevalent formant ratios in speech. Purves and his collaborators are now working on whether a given culture's preference for one subset of the tones over another is related to the formant relationships that are especially prevalent in the native language of that group.

Purves and his collaborators also think these findings may help explain a centuries-old debate in music over which tuning scheme for instruments works best. Ten of the 12 harmonic intervals identified in English and Mandarin speech occur in "just intonation" tuning, which sounds best to most trained musicians. They found fewer correspondences in other tuning systems, including the equal temperament tuning commonly used today.

Equal temperament tuning, in which each of the 12 interval distances in the chromatic scale is made exactly the same, is a scheme that enables an ensemble such as an orchestra to play together in different keys and across many octaves. Although equal temperament tuning sounds pretty good, it's a compromise on the more natural, vocally derived just intonation tuning system, Purves said.

The group's next study concerns our intuitive understanding that a musical piece tends to sound happy if it's in a major key but relatively sad if it's in a minor key. That, too, may come from the characteristics of the human voice, Purves suggests.

Source: Duke University

This document is subject to copyright. Apart from any fair dealing for the purpose of private study or research, no part may be reproduced without the written permission. The content is provided for information purposes only.