

Rough, smooth or deep: why the sound of a voice is multisensory

Pavlo Shopin is a lecturer in the department of philosophy of language, comparative linguistics and translation at the National Pedagogical Dragomanov University. He holds a PhD in German from the University of Cambridge.

Edited by Pam Weintraub

To make sense of human voices, we rely on senses beyond hearing. The songs of Taylor Swift can be sweet and soft. Lady Gaga's singing feels dark. Johnny Cash's voice was low and rough. That's because voice is not just sound: it can be seen and heard, but also tasted and touched. The sound we hear in voice creates 'multisensory images' – drawing in perceptions from many senses, not just one.

The phenomenon of multisensory perception can help us to understand why we assign metaphorical properties of softness, roughness or depth to voice. Think of a politician whose voice is flat. Flatness is a multisensory concept because it is both tactile and visual. We can recognise flat surfaces by either touching or seeing them. These sensory impressions inform us about the acoustic characteristics of voice, implying that it does not have variation in tone. Notably, flatness can also convey lack of sympathy and emotion on the part of the speaker.

Softness is another common way to present the auditory perception of sound. Like flatness, it can describe not only the sound quality but also the speaker's emotional state. And what about sharpness, a descriptor that might relate to both tactile and visual experience? Calling a voice sharp could be a metaphor for an aggressive, nasty speaker – or a means of describing acoustic, vocal sounds.

Multisensory images allow us to identify and deal with things that can harm or benefit us. A falling mortar shell, a jumping tiger or a skidding car are not just

auditory or visual images: they are perceived as multisensory images and can be conceived of as potential life threats. In cognitive psychology, it is generally recognised that, as Vanessa Harrar of the University of Oxford puts it, ‘integrating information from individual senses increases the chance of survival by reducing the variability in the incoming signals, thus allowing us to respond more rapidly’. In fact, notes Harrar, when the components of the multisensory signals are simultaneous, our reaction time is fastest of all.

The psychologist Charles Spence at the University of Oxford has done extensive research on how humans integrate sensory information with respect to culinary experience, finding that vision and hearing can change how food tastes. One study found that desserts tasted sweeter on a white plate than a black plate. Another study found that heavy cutlery made food taste better.

The multisensory perceptions that result in metaphors help us to think about relatively abstract things with more familiar ideas. In *Metaphors We Live By* (2003), the linguist George Lakoff and the philosopher Mark Johnson, who devised ‘conceptual metaphor theory’, say that humans use concrete ideas to understand abstract phenomena. Linguistic and psychological research supports the idea that metaphors empower our abstract thought about time, money, morality, death and even orgasm. Time, for example, is an abstract idea, and we tend to understand it through the more concrete-seeming experience of space: time can *flow*, and it can *stand still*. Our past is better *left behind*, because our future *lies ahead*.

Indifference or hostility are complex social concepts that can be conveyed through the experience of feeling cold. Coldness is tangible and vividly communicates the message. If someone’s voice is described as cold, people associate this sensory image with the emotional state of the speaker. In a similar vein, the acoustic properties of voice can be associated with other sensory experiences. A sharp voice can refer to both vision and the sense of touch.

Depicting how something ‘feels’ is one of the most common ways we use metaphors, especially when describing voices. That makes enormous sense because touch is a much earlier evolutionary development than speech, and is vital in daily life. In *Consciousness and the World* (2000), the Australian philosopher Brian O’Shaughnessy considered touch the primordial sense because ‘it is scarcely to be distinguished from the having of a body that can act in physical space’. And the evolutionary psychologist Robin Dunbar at the University of

Oxford argues that touch plays a significant role in social bonding in primates (including humans). Writing here in Aeon, the integrative biologist Steven Phelps at the University of Texas draws on Dunbar's research to argue that the use of touch for strengthening social relationships among primates appears to be 30 million years old.

Voice as a medium for language is a recent development in evolutionary terms, but it has become a crucial part of our social interactions. And it does not stand alone. We rely on a panoply of sensory experiences to navigate the medium of sound. The multisensory ensemble helps us to discuss a speaker's emotions and feelings through the conveyance of voice, creating interior meaning through metaphor. Description of touch and other senses can illuminate voice's deep meaning and its acoustic properties at once. Next time you hear a soft voice, reflect on the engaging feeling of softness that makes your experience so much more meaningful.