TONAL LANGUAGE PROCESSING

Fan-Gang Zeng
Departments of Anatomy and Neurobiology,
Biomedical Engineering, and Cognitive Sciences
University of California, Irvine
Irvine, California 92697

A tonal language uses changes in tone or pitch of a voiced sound to differentiate words. A classic example is the consonant-vowel combination /ma/ in Mandarin Chinese. The same /ma/, depending upon the tonal pattern of vowel /a/, can mean mother (妈, flat pattern), numb (麻, rising), horse (马, falling-rising), or curse (骂, falling). Growing up in the United States, my 9-year-old boy still confuses mother with horse, "cursing" his weekly 2-hour Chinese School as a form of "child abuse." Who should we blame for inventing tonal languages? What's good in them? Why is it hard for our brains? Or is it really?

> *"Why are 70% of the world's languages tonal and more than 2 billion people speaking them?"*

According to the late linguist, Yuen-Ren Chao, tones have been used to differentiate words in Chinese for at least 3,000 years. Recently researchers from the University of Edinburgh found that people who speak tonal languages also carry the least disturbed form of a 37,000-year-old gene, *Microcephalin*, suggesting that the first language was tonal (Dediu and Ladd, 2007). Indeed, ancient Greek (9-6th centuries BC) used tonal accents, but its tonality got lost, perhaps as a result of variations in that gene. Today, about 70% of the world's languages are tonal languages, which are spoken by over 2-billion people, mostly in sub-Saharan Africa and South East Asia (Haviland *et al.*, 2007). So, our ancestors out of Africa invented tonal languages, but why?

One answer may lie in the acoustics and perception of tones. Dr. Zhi-An Liang at the Shanghai Institute of Physiology published a classical paper (Liang, 1963) to show that compared with consonant and vowel perception, tone perception is the most redundant in terms of resiliency to acoustic distortions. Although tones are defined by variations in fundamental frequency, they can still be accurately perceived after removing the fundamental frequency via high-pass filtering or whispered speech. One can literally abuse the acoustic signal by filtering, infinite clipping, or adding noise, but still achieve a high level of tone perception. The reason for this high resistance to distortions and noise is that the acoustical cues for tone perception are multi-dimensional and widely distributed in both time and frequency domains. Tonal information is correlated with duration and temporal envelope in the time domain (Whalen and Xu, 1992; Fu *et al.*, 1998). But the more salient cues for tone perception are in the temporal fine structure, fundamental frequency, and their harmonics (Xu and Pfingst, 2003; Kong and Zeng, 2006). Possibly for their acoustical redundancy and perceptual resiliency, tones were invented to enable long distance communication in noisy backgrounds. Well, they are still used today by Spanish-speaking villagers who can whistle Silbo in the Canary Islands (Meyer, 2008) as well as tonal-language-speaking customers in a noisy Chinese restaurant (Lee, 2007; Luo *et al.*, 2009).

How do our ears and brain work together to process tonal information? Our ears are essentially filter banks that decompose sounds into different frequency regions. The filter bandwidth is narrow and relatively constant for center frequencies less than 2,000 Hz, but increases linearly for center frequencies above 2,000 Hz. In cases of a voiced sound, the fundamental frequency and its lower harmonics are likely separated into different filters, whereas the higher harmonics are likely combined into one filter. Tonal information is extracted from the output of these auditory filters.

There are at least three types of cues for pitch extraction. First, the fundamental frequency itself conveys a salient pitch percept by producing a strong timing cue that occurs in the right place or apical part of the cochlea. Second, the lower harmonics can also produce a salient pitch percept by generating a distinctive temporal and spatial pattern along the cochlea, a well-known phenomenon called the missing fundamental. Third, the unresolved high harmonics can produce a strong timing cue that is phase-locked to the fundamental frequency, but in the wrong place or basal part of the cochlea. Functionally, this envelope-based timing cue cannot provide a salient pitch percept (Zeng, 2002; Oxenham *et al.*, 2004).

Recent physiological studies have shed light on the brain's representation of pitch and its usage in tonal language processing. In marmoset monkeys, researchers found that neurons in a restricted low-frequency cortical region respond to both pure tones and their missing fundamental harmonic counterpart (Bendor and Wang, 2005). This cortical region has been mapped to Heschl's Gyrus in humans. Interestingly, in a study teaching English-speaking subjects to learn Mandarin tones, Wong and colleagues (2008) found that subjects, who were less successful in learning, showed a smaller Heschl's Gyrus volume on the left, but not on the right hemisphere, relative to learners who were successful. This finding leads to a general question on hemisphere specialization of tone perception: Which hemisphere do we use to process lexical tonal information?

Hemisphere specialization has been known for a long time in that the left hemisphere is for speech whereas the right hemisphere is for music processing. Tones are represented by changes in pitch—a salient music quality, but they also carry lexical meaning—a salient speech feature.

Recording brain waves from a group of Mandarin-speaking subjects, Luo and colleagues (2006) found a parsimonious answer in that hemisphere specialization of tonal processing is timing dependent. In early pre-attentive processing (<160 ms), like music, tonal information is processed by the right hemisphere, in which pitch information is presumably extracted. After that time mark and the extraction of pitch, tonal information encoding the lexical meaning is then processed, and guess where—in the left hemisphere. Combining the results from the Wong and Luo studies, we can infer that the failure for some English-speakers to learn Mandarin tones is not due to their inability to process pitch information in the right hemisphere, but rather their inefficiency to convert this pitch information into lexical meaning in the left hemisphere.

Processing deficits in both the ear and the brain contribute to tonal language perception and development by hearing-impaired listeners. In post-lingually hearing-impaired listeners, the deficit is mostly in the ear due to either poor spectral resolution as a result of hearing loss (Wang *et al.*, 2010), or lack of proper pitch extraction and delivery in auditory prostheses (Zeng *et al.*, 2008). For example, current cochlear implants do not extract, nor can they properly deliver, the salient pitch cue conveyed by the fundamental frequency or its lower harmonics. Instead, they only extract and deliver the less salient temporal envelope cue. As a result, tones are extremely difficult to produce and perceive by deaf children using a cochlear implant (Han *et al.*, 2007; Lee *et al.*, 2010). One solution to this problem is to deliver the fundamental frequency information acoustically via a hearing aid in the residual low-frequency region (<500 Hz) and simultaneously the high-frequency temporal envelope information via a cochlear implant. This "hybrid hearing" can improve tonal information transfer and speech perception in noise (Kong *et al.*, 2005; Qin and Oxenham, 2006). Surprisingly, introduction of the tonal information via hearing aids significantly helps cochlear-implant children to learn English, particularly the expressive part of the language (Nittrouer and Chapman, 2009). After all, tones are not just for tonal languages.**AT**

References

Bendor, D., and Wang, X. (**2005**). "The neuronal representation of pitch in primate auditory cortex." Nature **436**, 1161–1165.

Dediu, D., and Ladd, D. R. (**2007**). "Linguistic tone is related to the population frequency of the adaptive haplogroups of two brain size genes, ASPM and Microcephalin," Proceed. Natl. Acad. Sci. of the United States of America 104, 10944–10949.

Fu, Q. J., Zeng, F. G., Shannon, R. V., and Soli, S. D. (**1998**). "Importance of tonal envelope cues in Chinese speech recognition," J. Acoust. Soc. Am. **104**, 505–510.

Han, D., Zhou, N., Li, Y., Chen, X., Zhao, X., and Xu, L. (**2007**). "Tone production of Mandarin Chinese speaking children with cochlear implants," Intl. J. Pediatric Otorhinolaryngol. **71**, 875–880.

Haviland, W. A., Prins, H. E. L., Walrath, D., and McBride, B. (**2007**). *Anthropology: The Human Challenge* (12th ed.) (Thomson Higher Education, Belmont, CA).

Kong, Y. Y., and Zeng, F. G. (**2006**). "Temporal and spectral cues in Mandarin tone recognition" J. Acoust. Soc. Am. **120**, 2830–2840.

Kong, Y. Y., Stickney, G. S., and Zeng, F.-G. (**2005**). "Speech and melody recognition in binaurally combined acoustic and electric hearing," J. Acoust. Soc. Am. **117**, 1351–1361.

Lee, C. Y. (**2007**). "Does horse activate mother? Processing lexical tone in form priming," Language and Speech **50**, 101–123.

Lee, K. Y., van Hasselt, C. A., and Tong, M. C. (**2010**). "Lexical tone perception ability of profoundly hearing-impaired children: Performance of cochlear implant and hearing aid users," Otolol. and Neurotol. **31**, 1079–1087.

Liang, Z. A. (**1963**). "汉语普通话中声调的听觉辨认依据 The auditory discrimination basis of tone recognition in Standard Chinese," Acta Physiol. Sinica **26**, 85–91.

Luo, H., Ni, J. T., Li, Z. H., Li, X. O., Zhang, D. R., Zeng, F.-G., and Chen, L. (**2006**). "Opposite patterns of hemisphere dominance for early auditory processing of lexical tones and consonants." Proc. Natl. Acad. Sci. of the United States of America **103**, 19558–19563.

Luo, X., Fu, Q. J., Wu, H. P., and Hsu, C. J. (**2009**). "Concurrent-vowel and tone recognition by Mandarin-speaking cochlear implant users," Hearing Res. **256**, 75–84.

Meyer, J. (**2008**). "Typology and acoustic strategies of whistled languages: Phonetic comparison and perceptual cues of whistled vowels," J. Intl. Phonetic Assn. **38**, 69–94.

Nittrouer, S., and Chapman, C. (**2009**). "The effects of bilateral electric and bimodal electric–acoustic stimulation on language development," Trends in Amplification **13**, 190–205.

Oxenham, A. J., Bernstein, J. G., and Penagos, H. (**2004**). "Correct tonotopic representation is necessary for complex pitch perception," Proc. Natl. Acad. Sci. of the United States of America **101**, 1421–1425.

Qin, M. K., and Oxenham, A. J. (**2006**). "Effects of introducing unprocessed low-frequency information on the reception of envelope-vocoder processed speech," J. Acoust. Soc. Am. **119**, 2417–2426.

Wang, S., Mannell, R., Newall, P., and Han, D. (**2010**). "Contribution of spectral cues to Mandarin lexical tone recognition in normal-hearing and hearing-impaired Mandarin Chinese speakers," Ear and Hearing **32**, 97–103.

Whalen, D. H., and Xu, Y. (**1992**). "Information for Mandarin tones in the amplitude contour and in brief segments," Phonetica **49**, 25–47.

Wong, P. C., Warrier, C. M., Penhune, V. B., Roy, A. K., Sadehh, A., Parrish, T. B., and Zatorre, R. J. (**2008**). "Volume of left Heschl's Gyrus and linguistic pitch learning," Cerebral Cortex **18,** 828–836.

Xu, L., and Pfingst, B. E. (**2003**). "Relative importance of temporal envelope and fine structure in lexical-tone perception," J. Acoust. Soc. Am. **114**, 3024–3027.

Zeng, F.-G. (**2002**). "Temporal pitch in electric hearing," Hearing Res. **174**, 101–106.

Zeng, F.-G., Rebscher, S., Harrison, W., Sun, X., and Feng, H. H. (**2008**). "Cochlear implants: System design, integration and evaluation," IEEE Rev. in Biomed. Eng. **1**, 115–142.

Like father like son.

Fan-Gang Zeng is a leading researcher in auditory science and technology, unraveling brain mechanisms in loudness coding and speech recognition, while translating research into two commercial products for hearing loss and tinnitus treatment. He is a Professor of Anatomy and Neurobiology, Biomedical Engineering, Cognitive Sciences, and Otolaryngology, and Director of Center for Hearing Research at the University of California Irvine. He is a Fellow of the Acoustical Society of America, The American Institute for Medical and Biological Engineering, the Institute of Electrical and Electronics Engineers (IEEE), and the Collegium Oto-Rhino-Larygologicum.