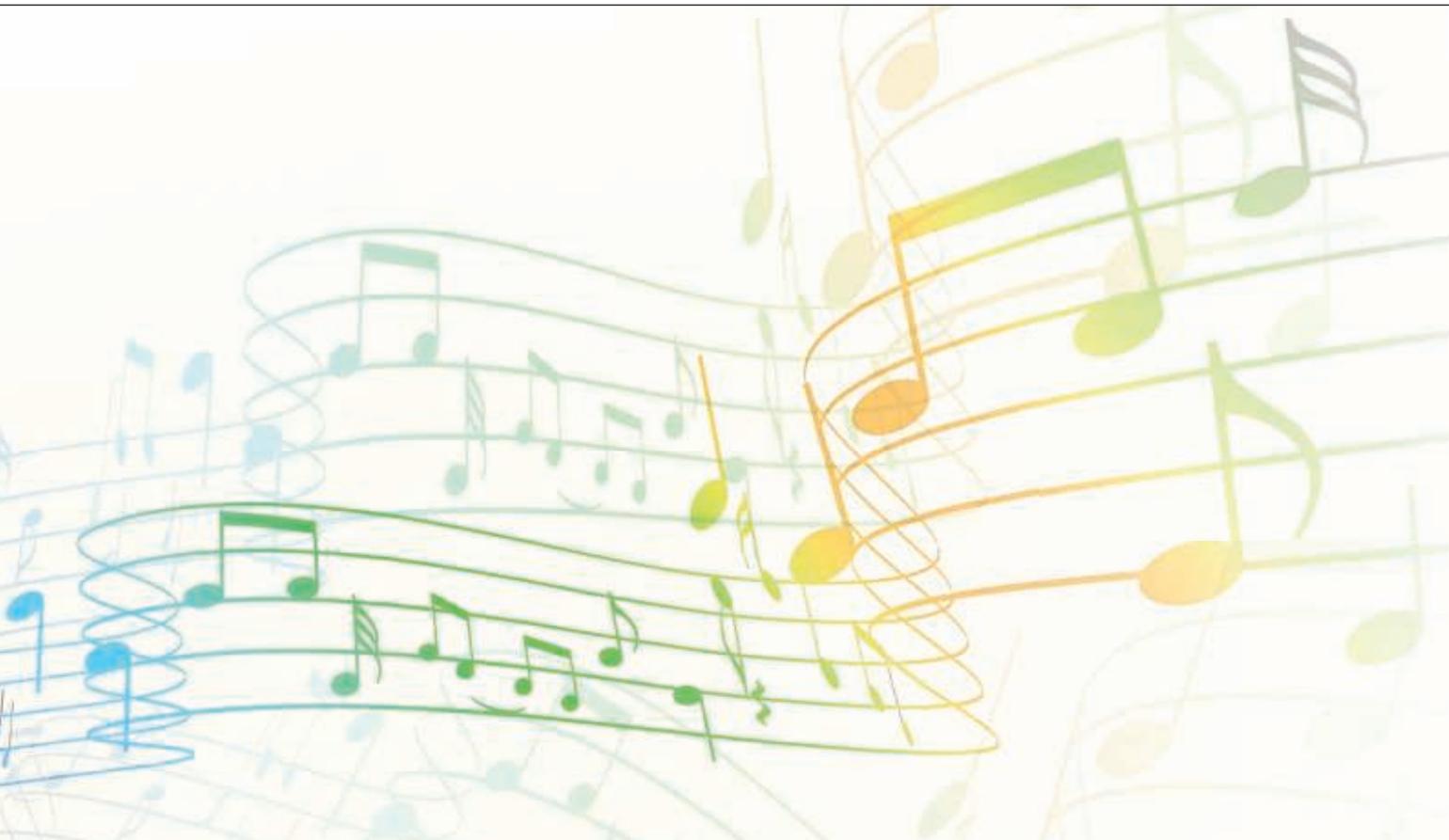


# Speaking in Tones

Music and language are partners in the brain. Our sense of song helps us learn to talk, read and even make friends

**By Diana Deutsch**



PIOTR POWIETRZYŃSKI age Fotostock (mouth); MARCELA BARSSE /Istockphoto (sheet music)

One afternoon in the summer of 1995, a curious incident occurred. I was fine-tuning my spoken commentary on a CD I was preparing about music and the brain. To detect glitches in the recording, I was looping phrases so that I could hear them over and over. At one point, when I was alone in the room, I put one of the phrases, “sometimes behave so strangely,” on a loop, began working on something else and forgot about it. Suddenly it seemed to me that a strange woman was singing! After glancing around and finding nobody there, I realized that I was hearing my own voice repetitively producing this phrase—but now, instead of hearing speech, I perceived a melody spilling out of the loudspeaker. My speech had morphed into song by the simple process of repetition.

This striking perceptual transformation, which I later found occurs for most people, shows that the boundary between speech and song can be very fragile. Composers have taken account of the strong connections between music and speech, for example, incorporating spoken words and phrases into their compositions. In addition, numerous vocalizations seem to fall near the boundary between speech and song, including religious chants and incantations, oratory, opera recitative (a style of delivery in

opera resembling sung ordinary speech), the cries of street vendors and some rap music.

And yet for decades the experience of musicians and the casual observer has clashed with scientific opinion, which has held that separate areas of the brain govern speech and music. Psychologists, linguists and neuroscientists have recently changed their tune, however, as sophisticated neuroimaging techniques have helped amass evidence that the brain areas governing music and language overlap. The latest data show that the two are in fact so intertwined that an awareness of music is critical to a baby’s language development and even helps to cement the bond between infant and mother. As children grow older, musical training may foster their communication skills and even their reading abilities, some studies suggest. What is more, the neurological ties between music and language go both ways: a person’s native tongue influences the way he or she perceives music. The same succession of notes may sound different depending on the language the listener learned growing up, and speakers of tonal languages such as Mandarin are much more likely than Westerners to have perfect pitch.

### Word Symphonies

Musicians and philosophers have long argued that speech and melody are interconnected. Russian



Certain vocalizations fall close to the border between speech and song. These include opera recitative, the shouts of street vendors and the chanting of rap musicians.

composer Modest Mussorgsky believed that music and talk were in essence so similar that a composer could reproduce a conversation. He wrote to his friend Rimsky-Korsakov: “Whatever speech I hear, no matter who is speaking ... my brain immediately sets to working out a musical exposition of this speech.” Indeed, when you listen to some of his piano and orchestral works, you may suddenly find that you are “hearing” the Russian language.

Despite such informal evidence of the ties between speech and music, researchers—bolstered in part by patients whose brain damage affected their speech but spared their musical ability—began espousing the opposite view around the middle of the 20th century. The brain divides into two hemispheres, and these experts hypothesized that its functions were just as neatly organized, with language residing on the left side and music on the right. Their theory was that the neural signal for dialogue bypassed the usual pathways for sound processing and instead was analyzed in an independent “module” in the brain’s left hemisphere. That module supposedly excluded nonverbal sounds such as music. Similarly, the theory went, music was processed in a right-hemisphere module that excluded

speech sounds. This attractive dichotomy became so popular that it effectively shut out for decades any thought that language and music might be neurologically—and functionally—intertwined.

But then, by the late 1990s, a generation of young researchers who did not have a stake in the separation of speech and song began questioning the idea. They brought to light existing data indicating that some aspects of music engage the left hemisphere more than the right. In addition, pioneering new experiments, many of which were conducted with emerging technology such as functional magnetic resonance imaging, showed that music and speech are not as neurologically separate as researchers had supposed.

One line of investigation demonstrated that the perception and appreciation of music could impinge on brain regions classically regarded as language processors. In a 2002 study neuroscientist Stefan Koelsch, then at the Max Planck Institute for Human Cognitive and Brain Sciences in Leipzig, Germany, and his colleagues presented participants with sequences of chords while using functional MRI to monitor their brains. They found that this task prompted activity on both sides of the brain but most notably in two regions in the left hemisphere, Broca’s and Wernicke’s areas [see illustration on opposite page], that are vital for language processing and that many researchers had assumed were solely dedicated to this function. Other more recent studies have revealed that speaking activates many of the same brain regions as analogous tasks that require singing. These and dozens of findings by other experimenters have established that the neural networks dedicated to speech and song significantly overlap.

This overlap makes sense, because language and music have a lot in common. They are both governed by a grammar, in which basic elements are organized hierarchically into sequences according to established rules. In language, words combine to form phrases, which join to form larger phrases, which in

#### FAST FACTS

### Singing in the Brain

- 1>> The brain circuits that interpret music overlap with those that process speech.
- 2>> The musical qualities of speech are critical in early language development and in cementing the bond between infant and mother.
- 3>> A person’s native language may affect the way he or she hears a set of musical notes.

AFP/GETTY IMAGES (opera singers); GETTY IMAGES (vendor); NEWSPOPT/CORBIS (rapper)

# Language and music are both governed by a grammar: basic elements organized hierarchically into sequences.

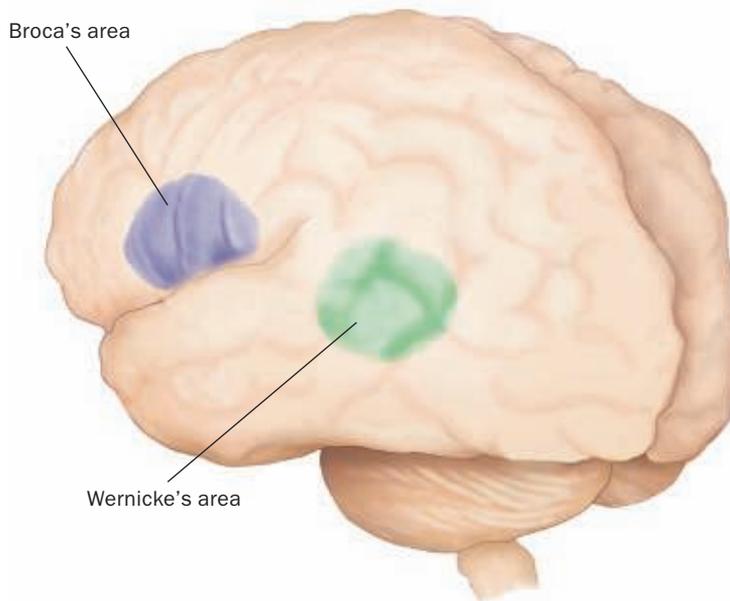
turn combine to make sentences. Similarly, in music, notes combine to form phrases, which connect to form larger phrases, and so on. Thus, to understand either language or music, listeners must infer the structure of the passages that they hear, using rules they have assimilated through experience.

In addition, speech has a natural melody called prosody. Prosody encompasses overall pitch level and pitch range, pitch contour (the pattern of rises and falls in pitch), loudness variation, rhythm and tempo. Prosodic characteristics often reflect the speaker's emotional state. When people are happy or excited, they frequently speak more rapidly, at higher pitches and in wider pitch ranges; when people are sad, they tend to talk more slowly, in a lower voice and with less pitch variation. Prosody also helps us to understand the flow and meaning of speech. Boundaries between phrases are generally marked by pauses, and the endings of phrases tend to be distinguished by lower pitches and slower speech. Moreover, important words are often spoken at higher pitches. Interestingly, some pitch and timing characteristics of spoken language also occur in music, which indicates that overlapping neural circuitries may be involved.

## Meaningful Melodies

At birth, babies are already familiar with the melody of their mother's speech. Audio recordings taken from inside the womb at the beginning of labor reveal that speech sounds produced by the mother can be loudly heard. The phrases reaching the baby have been filtered through the mother's tissues, however, so that the crisp, high frequencies—which carry much of the information important for identifying the meanings of words—are muted, whereas the musical characteristics of speech—its pitch contours, loudness variations, tempo and rhythmic patterning—are well preserved.

These spoken melodies seem to set the stage for mother-child bonding. In an ingenious experiment published in 1980, psychologists Anthony J. DeCasper of the University of North Carolina at Greensboro and William P. Fifer, now at Columbia University, recorded new mothers reading a story out loud. In this experimental setup, the newborn babies could turn on the recordings by sucking on a pacifier, a connection they learned over time, and they sucked more frequently when their actions produced their mothers' voices compared with those of other women. The researchers reasoned that the newborns preferred to



listen to the voices with which they had become familiar before birth. Then, in 1996, psychologists Melanie J. Spence and Mark S. Freeman of the University of Texas at Dallas reported carrying out a similar experiment in which they used a low-pass filter to muffle recorded female voices so that they sounded as they would in the womb. The newborn babies preferred their mothers' filtered voices over those of other women, again indicating that they had become familiar with the melodies of their mothers' utterances in the womb.

In addition to forging a nascent connection between mother and child, early exposure to musical speech sounds may begin the process of learning to talk. In one 1993 study, for example, two-day-old babies preferred to listen to recordings of speech in their native language to those in a foreign tongue. Because such young babies could only have become familiar with such speech in the womb, the results suggest that the babies initially become comfort-

The appreciation of music requires many of the same brain regions that are involved in the processing of language. These multipurpose regions include Broca's (purple) and Wernicke's (green) areas.

## (The Author)

DIANA DEUTSCH is a professor of psychology at the University of California, San Diego, who studies the perception of music and language. She has recorded two CDs consisting of audio illusions that she has discovered: *Musical Illusions and Paradoxes* and *Phantom Words and Other Curiosities* (<http://philomel.com>). In these anomalies, the perception of a given set of sounds varies among people or changes over time, even when the sounds remain the same. For examples, see <http://deutsch.ucsd.edu>.

The exaggerated speech melodies—termed motherese—that parents use when speaking to their babies help the infants grasp the speaker's intentions.



speech patterns termed motherese that are characterized by high pitches, large pitch ranges, slow tempo, long pauses and short phrases. These melodious exaggerations help babies who cannot yet comprehend word meanings grasp their mothers' intentions. For example, mothers use falling pitch contours to soothe a distressed baby and rising pitch contours to attract the baby's attention. To express approval or praise, they utter steep rising and falling pitch contours, as in "Go-o-o-d girl!" When they express disapproval, as in "Don't do that!" they speak in a low, staccato voice.

In 1993 psychologist Anne Fernald of Stanford University reported exposing five-month-old infants from English-speaking families to approval and prohibition phrases spoken in German, Italian and non-sense English, as well as regular English motherese. Even though all this speech was gibberish to the babies, they responded with the appropriate emotion, smiling when they heard approvals and becoming subdued or crying when they heard prohibitions. Thus, the melody of the speech alone, apart from any content, conveys the message.

Although the ability to detect speech melodies

## Infants echo the inherent melodies of their native language when they cry, long before they can speak.

able with the musical qualities of their language.

Accordingly, music may be the first part of speech that babies learn to reproduce; infants echo the inherent melodies of their native language when they cry, long before they can utter actual words. In a study published in 2009 medical anthropologist Kathleen Wermke of the University of Würzburg in Germany and her colleagues recorded the wails of newborn babies—which first rise and then fall in pitch—who had been born into either French- or German-speaking families. The researchers found that the cries of the French babies consisted mostly of the rising portion, whereas the descending segment predominated in the German babies' cries. Rising pitches are particularly common in French speech, whereas falling pitches predominate in German. So the newborns in this study were incorporating into their cries some of the musical elements of the speech to which they had been exposed in the womb, showing that they had already learned to use some of the characteristics of their first language.

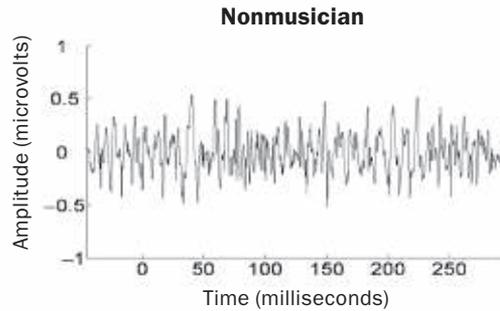
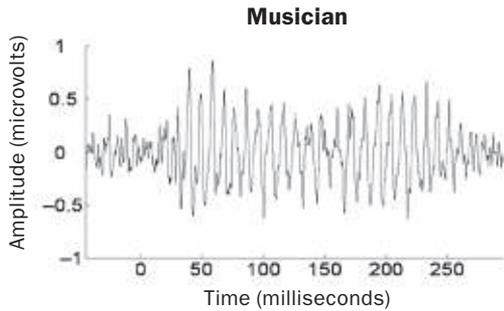
After birth, the melody of speech is also vital to communication between mother and infant. When parents speak to their babies, they use exaggerated

is inborn, people can hone this skill by taking music lessons. In a study published in 2009 neuroscientists Mireille Besson of CNRS in France and Sylvain Moreno, now at the Rotman Research Institute in Toronto, and their colleagues recruited eight-year-old children who had been given no musical training and divided them into two groups. One group took music lessons for six months while the other enrolled in painting lessons. Before and after this training, the children listened to recorded sentences; in some of these, the last word was raised in pitch so that it sounded out of keeping with the rest of the sentence, and the children were asked to detect the altered sentences. At the start, the two groups did not differ in their ability to detect the pitch changes, but after the six months of instruction, the children who had taken music lessons outperformed the others. Musically trained children may thus be at an advantage in grasping the emotional content—and meaning—of speech.

Musical training may affect perception of prosody in part by tuning the auditory brain stem—a group of structures that receive signals from the ear and help to decode the sounds of both speech and

# Tuning the Brain

In a 2007 study scientists monitored the brain waves of people who were listening to recordings of words spoken in Mandarin. The study subjects, all English speakers, did not understand the meaning of what they heard. But the ones who had a music background showed a much stronger electrical response in a region called the auditory brain stem (*left*) than did those who had no musical training (*right*). The work suggests that learning to sing or play an instrument can make people more attuned to the melody of speech.



music. In a 2007 investigation neuroscientists Patrick Wong and Nina Kraus, along with their colleagues at Northwestern University, exposed English speakers to Mandarin speech sounds and measured the electrical responses in the auditory brain stem using electrodes placed on the scalp. The responses to Mandarin were stronger among participants who had received musical training—and the earlier they had begun training and the longer they had continued training, the stronger the activity in these brain areas [see illustration above].

Additional research shows that music lessons can improve the ability to detect emotions conveyed in speech (presumably through a heightened awareness of prosody). In a study published in 2004 psychologist William F. Thompson and his colleagues at the University of Toronto gave a group of six-year-old children musical keyboard lessons for a year and then tested their ability to identify emotions expressed in spoken sentences, comparing their scores with those of children who did not receive musical training. They found that the kids who received music lessons were better at identifying whether sentences were spoken in a fearful or angry tone of voice—even when the sentences were spoken in an unfamiliar language.

Musical training might even accelerate the process of learning to read. Good readers tend to do better than poor readers on tests of musical ability (although there are many exceptions to this rule). In their 2009 study Moreno and his colleagues found that the eight-year-olds who had taken music lessons also showed better reading ability than the children who had instead learned to paint, suggesting that facility with music may spill over into skill at deciphering the written word. Researchers have even suggest-

ed that musical training (in combination with other therapies) might be useful in remedying dyslexia.

## Talking in Tune

Not only can exposure to music enhance our language skills, but the speech we hear also influences our perception of music. For example, in a musical illusion called the tritone paradox, which I discovered in the 1980s, a listener hears two computer-generated tones that are half an octave (or tritone) apart, one after the other. Each tone is a clearly defined note such as C, C-sharp or D, but its octave is inherently ambiguous so that a note could be, say, middle C, an octave above or below middle C, or any other C. The listener then decides whether the pattern ascends or descends in pitch. (Because of the ambiguity in the octave placement of the notes, there is no correct answer, and perception varies by listen-

Taking music lessons can help children detect emotions conveyed in speech and may even accelerate the process of learning to read.



➔ For an audio slideshow featuring Deutsch's work on the musical aspects of language, visit [www.ScientificAmerican.com/Mind/music-and-speech](http://www.ScientificAmerican.com/Mind/music-and-speech)

er.) Interestingly, I found that such judgments depend on the language or dialect to which the listener has been exposed. For example, in a 1991 study I asked people who had been raised in California and those raised in the south of England to judge these tritones and found that when the Californians tended to hear the pattern as ascending, the southern English subjects tended to hear it as descending, and vice versa. In another study published in 2004 my colleagues and I found the same dichotomy between listeners from Vietnam and native English speakers born in California, suggesting that the language we learn early in life provides a *musical* template that influences our perception of pitch.

Such a template might also constrain the pitch range of our speaking voice. In a study published in 2009 my colleagues and I examined the pitch ranges of female speech in two Chinese villages and found that these clustered together for people in the same village but differed across villages, suggesting that even local differences in the voices we hear around us can affect the pitch of the speech we produce.

The language to which we are exposed can also greatly influence the chances of developing perfect pitch—the ability to name the pitch of a note without a reference note. This skill is very rare in our culture: only an estimated one in 10,000 Americans have it. In 1997 I noticed that when I uttered a Vietnamese word without paying attention to its pitch, a native

listener would either misunderstand me or have no idea what I was trying to say. But when I got the pitch right, the problem disappeared. Vietnamese and Mandarin are tone languages in which words take on entirely different meanings depending on the tones with which they are spoken. In Vietnamese, the word “ba” spoken in the mid-flat tone means “father;” the same word spoken in the low-descending tone means “grandmother.” In Mandarin, the word “ma” means “mother” in a tone that is high and flat but “horse” in a tone that is low and first descends and then ascends [see table on opposite page].

I then learned that not only were Vietnamese and Mandarin speakers very sensitive to the pitches that they hear, but they can produce words at a consistent absolute pitch. In a study published in 2004 my colleagues and I asked native speakers of Mandarin and Vietnamese to recite a list of words in their native language on two separate days. We found that their pitches were remarkably consistent: when compared across days, half of the participants showed pitch differences of less than half a semitone. (A semitone is half a tone—that is, the difference between F and F-sharp.)

In light of these findings, I wondered if tone language speakers acquire perfect pitch for the tones of their language in infancy along with other features of their native tongue. Perfect pitch for musical tones would then be much easier for tone language

## The language we learn early in life provides a musical template that influences our perception of pitch.

Perfect pitch is remarkably common in speakers of tone languages. Ninety-two percent of Mandarin speakers who began music lessons at or before age five had perfect pitch, compared to just 8 percent of English speakers with comparable musical training.



HANS NELEMAN Getty Images

# What's in a Word?

In a tone language such as Mandarin, a word's meaning depends on its pitch and pitch pattern. Below are examples of words that are phonetically identical but, depending on the pitches with which they are spoken, have four unrelated meanings. (Tone 1 is high and flat; tone 2 starts mid-high and rises; tone 3 starts low, falls and then rises; and tone 4 begins high and falls.) Each word is written with a distinct character.

Tone	Word	Chinese character	English meaning
1	mā	妈	mother
2	má	麻	hemp
3	mǎ	马	horse
4	mà	骂	reproach

Tone	Word	Chinese character	English meaning
1	wēn	温	warm
2	wén	闻	hear
3	wěn	稳	stable
4	wèn	问	ask

Tone	Word	Chinese character	English meaning
1	yāo	腰	waist
2	yáo	摇	shake
3	yǎo	咬	bite
4	yào	药	drug

Tone	Word	Chinese character	English meaning
1	wēi	微	tiny
2	wéi	围	surround
3	wěi	尾	tail
4	wèi	喂	feed

speakers to develop than it would be for speakers of a nontone language, such as English. In an experiment published in 2006 my colleagues and I gave a test for perfect pitch to two large groups of music conservatory students—Mandarin speakers at the Central Conservatory of Music in Beijing and speakers of English or of another nontone language at Eastman School of Music in Rochester, N.Y.—and found that the prevalence of perfect pitch was indeed far higher among the Mandarin speakers. These findings were consistent with my hypothesis, but because the Central Conservatory students were all Chinese, the results could mean that genes that spur the development of perfect pitch are just more prevalent among Chinese people.

To decide which explanation was correct, my colleagues and I gave a test for perfect pitch to University of Southern California music conservatory students, including English speakers and three groups of East Asian students divided by how well they spoke their native tone language. Among the English speakers, the prevalence of perfect pitch was just 8 percent among those who had begun musical training at or before age five and 1 percent among those who had begun training between ages six and nine. The statistics were similar among the East Asian students who were not at all fluent in their native tone language. In contrast, the students who were very fluent tone language speakers performed extraordinarily well on our test: 92 percent

of those who had begun musical training at or before age five had perfect pitch as did 67 percent of those who started music lessons between ages six and nine. The students who spoke a tone language moderately well fell between the two extremes. These findings, which we published in 2009, strongly indicate that the high prevalence of perfect pitch among the tone language speakers is not genetic but related to exposure to their language.

Thus, the language we learn in infancy, and continue to speak, can have a profound effect on the way in which we encode the sounds of music. Indeed, in many respects, music and speech seem to be mirror images, with both playing integral roles in the development of the other—in the way we, as people, bond and communicate, in how we perceive the sounds around us, in our understanding of language and in the workings of our minds. **M**

## (Further Reading)

- ◆ **The Psychology of Music.** Second edition. Edited by Diana Deutsch. Elsevier, 1999.
- ◆ **The Enigma of Absolute Pitch.** Diana Deutsch in *Acoustics Today*, Vol. 2, pages 11-19; 2006.
- ◆ **Musicophilia: Tales of Music and the Brain.** Oliver Sacks. Knopf, 2007.
- ◆ **Newborns' Cry Melody Is Shaped by Their Native Language.** Birgit Mampe et al. in *Current Biology*, Vol. 19, pages 1994-1997; 2009.
- ◆ **Perfect Pitch: Language Wins Out over Genetics.** Diana Deutsch et al.: [www.acoustics.org/press/157th/deutsch.html](http://www.acoustics.org/press/157th/deutsch.html)
- ◆ **The Speech-to-Song Illusion.** Diana Deutsch et al.: [www.acoustics.org/press/156th/deutsch.html](http://www.acoustics.org/press/156th/deutsch.html)