In search of the language of music

Victoria Williamson compares two human universals

Comparing language and music is a tradition inherited from Aristotle and Darwin. The aim of the game is to learn more about how we interact with our auditory environment by understanding why language and music evolved, how we learn about them as children, how we use them as adults and how they are processed in the brain.

This article discusses the nature of language and music as well as a selection of the current directions research has taken in order to compare them. The hunt is on to discover the similarities and differences between these two universal forms of communication.

resources

Patel, A.D. (2009). Music and the brain: Three links to language. In S. Hallam, I. Cross & M. Thaut (Eds) *The Oxford handbook of music psychology* (pp. 208–216). Oxford: Oxford University Press.
Jackendoff, R. (2009) Parallels and nonparallels between language and music. *Music Perception, 26*, 195–204.
www.youtube.com/watch?v=2oMvtw4aeEY (lectures in the music in the brain series)

Music expresses that which cannot be said and on which it is impossible to be silent.

In the above quotation Victor Hugo captures a widely acknowledged attribute of music – the sense that it helps us to communicate. It has been described, for example, as the language of love ('Music is love in search of a word,' said American poet Sidney Lanier). But aside from romantic aspirations, how far can we really take the analogy between language and music? Did they evolve together or separately? What are the similarities between their structures? What can the relationship between them tell us about the way our brains are organised? How might knowledge of music influence language learning?

These types of question have fuelled interest and research since the time of Aristotle. After all, language and music are both human universals. There is not now, nor has there ever been, a known human society that does not use language and music in some form (Wallin et al., 2000). The answers to the questions above have proved surprising at times, but this area of study has the potential to elucidate how and why our lives are so intricately interwoven with these two sound systems.

What are language and music?

Language is the principle form of human communication. Although a number of the world's languages are silent (e.g. sign or gesture), most of the research in this area has compared spoken language and music. Speech consists of sounds that we create by moving our mouth, tongue and lips in order to produce phonemes and syllables. Combinations of these sounds are then assembled, using the rules of grammar, to form words and sentences that represent our thoughts, actions and feelings.

A definition of music is a little harder to pin down. At the basic level music is a collection of sounds produced by the human voice or separate instrument . Beyond this there are very few universals that apply to musical form (Nettl, 2000), but the majority of research in music psychology has focused on tonal music (a tradition of scale structure and harmony that emerged in European music during the Renaissance).

So both speech and music are complex auditory sequences that unfold over time.

references

Besson, M., Schön, D., Moreno, S. et al. (2007). Influence of musical expertise and musical training on pitch processing in music and language. *Restorative Neurology and Neuroscience, 25*, 399–410.
Brown, S., Martinez, M.J., Hodges, D.A. et al. (2004). The song system of the human mind. *Cognitive Brain Research, 20*, 363–375.
Cross, I. (1999). Is music the most important thing we ever did? In S.W. Yi (Ed.) *Music, mind and science* (pp.10–39). Seoul: Seoul National University Press.
Darwin, C. (1871). *The descent of man, and selection in relation to sex.* New York: D. Appleton and Company.
Fedorenko, E., Patel, A.D., Casasanto, D. et al. (2009). Structural integration in language and music. *Memory and Cognition, 37*(1), 1–9.
Halpern, A.R. & Zatorre, R.J. (1999). When the tune runs through your head: A PET investigation of auditory imagery for familiar melodies. *Cerebral Cortex, 9*, 697–704.
Ho, Y-C., Cheung, M.C. & Chan, A.S. (2003). Musical training improves verbal but not visual memory. *Neuropsychology, 17*, 439–450.
Koelsch, S., Schultz, K., Sammler, D. et al. (2008). Functional architecture of verbal and tonal working memory. *Human Brain Mapping, 30*(3), 859–873.
Koelsch, S., Gunter, T.C., Wittforth, M. et al. (2005). Interaction between syntax processing in music and language: An ERP study. *Journal of Cognitive Neuroscience, 17*, 1565–1577.
McMullen, E. & Saffran, J.R. (2004). Music and language: A developmental comparison. *Music Perception, 21*(3), 289–311.

But can we compare them beyond this elementary level? Let's go back to the beginning, to theories about how language and music evolved.

The singing Neanderthals

In his book *The Singing Neanderthals*, Steven Mithen (2005) presented a modern version of an argument that can trace its origins to Darwin. In *The Descent of Man, and Selection in Relation to Sex* (1871), Darwin postulated that the development of musical behaviours in animals and humans was driven by the pressures of sexual selection. He suggested that at some point the musical sounds humans produced with the intention to attract mates began to serve as a 'protolanguage'. The implication is that at this intermediate stage of development there was the potential for commonality between the generation and production of the sounds that we now define as language and music.

Steven Pinker (1997) rejects the idea of a musical protolanguage. He most controversially described music, given its apparent lack of adaptive function in human evolution, as 'auditory cheesecake'. He argued that music was essentially a lucky by-product of the evolution of language. Like many people's love of cheesecake, he suggested that love of music had emerged because we developed the sensory and cognitive systems that responded to their components (fats and sugars in the one case and complex patterns of sounds in the other), not because we had a particular need or liking for them.

Although these two opposing viewpoints occupy seemingly distant ideological camps, both postulate a single precursor in history, whether it was language or a more music-based form of communication that came first. So if there was once similarity in the cognitive systems designed to process language and music, there may still be some overlap today. Of course there is no way to further argue this point from the evolutionary evidence alone. Whilst we have early controversial examples of musical instruments and cave writings the vast majority of the physical evidence regarding the development of auditory music and language has been lost in the mists of time.

Follow the rules

One way to compare the similarities in modern speech and music is to examine their structures. Speech and music are constructed across multiple auditory dimensions that unfold over time (pitch, timbre, rhythm). Because of their physical nature they are both initially treated in similar ways by our auditory systems. The airborne vibrations are transformed into a physical signal in the middle ear. This signal is then converted into a neural signal by the action of the structures within the cochlear. But what happens once these neural signals get to the brain?

The first job is to assemble meaningful streams of sound. A major similarity between language and music is their 'particulate' or 'combinatory' nature. This means that they are both made of small elements (notes or syllables) that can be assembled in a seemingly infinite number of combinations (melodies or sentences) using hierarchical structural rules or syntax (Patel, 2008). In language these rules are termed grammar and in music they can be tonality, harmony, or form.

In 2003 Aniruddh Patel proposed the shared syntactic integration resource hypothesis (SSIRH). The SSIRH argues that whilst representations of language and music are likely to be stored independently in the brain, we use similar neural networks to integrate evolving speech and music sounds. In essence Patel suggests that whilst the ingredients for language and music may be different, the way we build structure from them is very similar. Evidence that has tested predictions of the SSIRH has so far been very supportive (Fedorenko et al., 2009; Koelsch et al., 2005; Slevc et al., 2007).

What happens if we fail to build this structure? We all instinctively recognise the consequences for disobeying the rules of grammar. We may be misunderstood or fail to get our message across entirely. In music the consequences may not seem so obvious, but our preference for certain musical structures explains in large part why we find it difficult to listen to music from other cultures, where different rules apply (e.g. Javanese gamelan or Japanese gagaku). It also explains why many people find it difficult to listen to atonal music, where the rule book is figuratively thrown out the window.

Violations of expectancy can also impact on the meaning we interpret from music and language (Steinbeis et al., 2007). Steinbeis and Koelsch (2008) used an affective priming paradigm to investigate whether music meaning could influence subsequent processing of language meaning, and vice versa. In other words, if the music/language you hear leads you to expect a pleasant sound in the other domain (e.g. the word 'love' or a consonant chord), does your processing slow down when you hear an unpleasant sound (e.g. the word 'hate' or a dissonant chord)? The authors found that people's responses were faster when the prime matched the target. ERP data also showed similar brain wave patterns associated with incongruous pairings; suggesting that meaning of language and music may trigger comparable responses in the brain. The authors argued against a generalised emotion explanation for the similarities; the brain activity they identified occurred outside those areas traditionally associated with emotion conflict. It should be noted, however, that the fMRI data they collected suggested that the processing of language

> "how far can we really take the analogy between language and music?"

Mithen, S. (2005). *The singing Neanderthals: The origins of music, language, mind and body.* London: Weidenfield & Nicholson.

Nettl, B. (2000). An ethnomusicologist contemplates universals in musical sound and musical culture. In N.L. Wallin, B. Merker & S. Brown (Eds.) *The origins of music* (pp.463–472). Cambridge, MA: MIT Press.

Norton, A., Zipse, L., Marchina, S. &

Schlaug, G. (2009) Melodic Intonation Therapy. In The Neurosciences and Music III: Disorders and Plasticity. *Annals of the New York Academy of Sciences, 1169,* 431-436.

Patel, A.D. (2003). Language, music, syntax, and the brain. *Nature Neuroscience, 6*(7), 674–681.

Patel, A.D. (2008). *Music, language, and the brain.* New York: Oxford University Press.

Peretz, I. & Zatorre, R.J. (2005). Brain organisation for music processing. *Annual Review of Psychology, 56,* 89–114.

Pinker, S. (1997). *How the mind works.* London: Allen Lane.

Schellenberg, E.G. (2004). Music lessons enhance IQ. *Psychological Science, 15,* 511–514.

Slevc, L.R., Rosenberg, J.C. & Patel, A.D. (2007). Language, music, and modularity. Presentation at the 20th

CUNY Sentence Processing Conference, San Diego, CA.

Steinbeis, N. & Koelsch, S. (2008). Comparing the processing of music and language meaning using EEG and fMRI provides evidence for similar and distinct neural representations. *PLoS ONE, 3*(5), e2226. doi: 10.1371/journal.pone.0002226.

Steinbeis, N., Koelsch, S. & Sloboda, J. (2007). The role of harmonic

and music meaning did not occur in identical neural structures.

Baby talk is music

Recent research has revealed a number of striking similarities in the way that children come to learn about language and music. Infant directed speech (IDS) is often referred to as musical speech because of its exaggerated use of pitch and contour (the patterns of ups and downs). Interestingly the features of IDS seem to be fairly universal despite the variations in the languages of the world (Trainor & Desjardins, 2002).

The purpose of IDS is still debated, but some theories suggest that the musical aspects help to communicate emotion.

Does music hinder or help?

We often try to combine language and music processing (e.g. listening to iPod while studying) but do these types of task result in a drain on shared resources? Even if you rule out effects of auditory distraction, evidence suggests that music can interfere with performance on a language task, although typically not as much as background speech. But the news is not all bad. Some of the negative effects of music in a dual task environment result from altering mood and/or arousal levels, and this can have cognitive benefits for some individuals.

Furthermore, it appears that music can actually help recover lost language abilities. Melodic intonation therapy encourages patients with conditions such as aphasia to capitalise on preserved singing function and regain expressive language communication by utilising the inherent melody in speech. Over the course of treatment, patients can enhance fluency and improve conversational skills (Norton et al., 2009). Although this is an exciting application of comparative language/music research, we do not yet fully understand the underlying mechanisms.

Trainor et al. (2000) analysed the different acoustic attributes of IDS and emotional adult speech: pitch range, contours, tempo and rhythm. They reported that the pitch of IDS tended to be higher but otherwise there was little difference on the other measures taken. This might suggest that musical sounds are important in the way that children learn about the overall emotion and intention of speech before knowledge of direct meaning develops.

Another hypothesis is that the musical aspects of IDS are directly involved in how children learn language. Trainor and Desjardins (2002) looked at the effect of IDS on six- to seven-month-old children's ability to discriminate between vowel sounds. They recorded two English vowels with flat or moving pitch contours (as used in IDS), and measured how often babies turned their head towards the speaker when they changed the stimulus (a method known as the conditioned head turn procedure). They found correct responses were significantly greater when a moving pitch contour was presented as opposed to a flat contour. The authors suggested that the musical aspects of IDS may play a role in helping babies to identify important syllable and word boundaries in speech.

McMullen and Saffran (2004) have argued that there is an aspect of statistical learning to both language and music that necessarily suggests a high level of overlap in the underlying processes that may be involved. This may explain why babies show little preference for either their native language or their culture's musical conventions early in life (Trehub et al., 1999). At this stage babies are still gathering information about these sounds and internalising the patterns and structures that they hear regularly. It is only later that they develop a preference for the sounds that they are used to hearing, be they speech or music. This exciting field of research is

set to continue, hopefully testing in multiple languages and musical traditions.

The case of memory

Neuroimaging can tell us about the way that language and music are processed in the brains of adults. A lot of research suggests that music is processed in distinct areas of the brain separate from those used to process language (Peretz & Zatorre, 2005). However, a number of recent studies have identified common activation for tasks that involve language and music memory (Koelsch et al., 2008).

To give a specific example, Brown et al. (2004) conducted a positron emission tomography (PET) study of melody imitation from memory. The authors found activation in Broca's area and the supplementary motor area. Both these areas of the brain have also been shown to

expectancy violations in musical emotions. *Journal of Cognitive Neuroscience, 18*, 1380–1393.

Tanaka, A. & Nakamura, K. (2004). Auditory memory and proficiency of second language learning. *Psychological Reports, 95*, 723–734.

Trainor, L.J., Austin, C.M. & Desjardins, R.N. (2000). Is infant-directed speech prosody a result of the vocal expression of emotion? *Psychological*

Science, 11(3), 188–195.

Trainor, L.J. & Desjardins, R.N. (2002). Pitch characteristics of infant-directed speech affect infants' ability to discriminate vowels. *Psychological Bulletin and Review, 9*(2), 335–340.

Trehub, S.E., Schellenberg, E.G. & Kamenetsky, S.B. (1999). Infants' and adults' perception of scale structure. *Journal of Experimental Psychology: Human Perception and Performance,*

25, 965–975.

Vallar, G. (2006). Memory systems: The case for phonological short-term memory. *Cognitive Neuropsychology, 23*(1), 135–155.

Vines, B.W., Schnider, N.M. & Schlaug, G. (2003). Testing for causality with transcranial direct current stimulation: Pitch memory and the left supramarginal gyrus. *Cognitive Neuroscience and Neuropsychology,*

17(10), 1047–1050.

Williamson, V.J., Baddeley, A.D. & Hitch, G.J. (in press). Musicians' and nonmusicians' short-term memory for verbal and musical sequences. *Memory and Cognition*.

Wong, P.C.M., Skoe, E., Russo, N.M. et al. (2007). Musical experience shapes human brainstem encoding of linguistic pitch patterns. *Nature Neuroscience, 10*, 420–422.

sounds as measured both in behavioural tests and in auditory pathway responses (Besson et al., 2007; Wong et al., 2007).

Whether these improvements represent a general benefit to auditory processing of all sounds or one that is specific to language has yet to be firmly established. It may also be the case that people with inherent advantages in auditory processing are more likely to take up and maintain musical training. If this is true, then the differences we are detecting may exist before the musical training is introduced. Longitudinal studies are needed of children's abilities before and after training to confirm or refute this confound. These studies are being carried out now and I for one am eagerly anticipating the results!

Conclusions

Despite the apparent similarities there are fundamental differences between language and music that must not be forgotten. One point is that language is a referential form of communication; we use language to refer to objects in the real world. If I say 'dog' we can picture the word itself, as well as perhaps images and knowledge of dogs. However, music is largely a-referential. To imagine a piece of music that everyone would recognise as depicting a dog is very difficult indeed. The 'floating intentionality' (Cross, 1999) of music is at the heart of its ability to engage different emotions and meanings in different people. A composer once told me that great music is built like a mirror, in which all listeners see themselves.

This is not the only point of divergence between language and music. The point is that a research-style health warning should be attached to the work in this field: Music is not a language. There are many fundamental differences, but drawing out similarities makes it clear that there is value in the comparative study of language and music. By studying their interactions and overlaps, we can learn far more about the way we engage with our auditory environment than we ever can by studying either domain in isolation. Continuing to explore how and why we come to employ language and music will add another important dimension in our quest to learn more about what it means to be human.

I **Victoria Williamson**
is a Postdoctoral Fellow at Goldsmiths, University of London
vickywilliamson01@googlemail.com

be active when people are silently rehearsing music (Halpern & Zatorre, 1999) and speech (Vallar, 2006).

A valid criticism is that neuroimaging shows no direct causal relationship between the stimulus being presented and the subsequent brain activation. One way to address this issue is to use transcranial direct current stimulation (tDCS). This involves placing two electrodes (one positively and one negatively charged) on the surface of the scalp. A small current is then passed through the electrodes that can penetrate the skull. This can reduce the neural firing rate in the affected area, effectively creating a temporary virtual lesion. Vines et al. (2003) used this technique to target an area of the brain associated with language memory. They then tested participants' memory for pitch. They found that scores were poorer on the pitch memory task immediately after the tDCS was applied. The authors argued this was evidence that at least one area of the brain associated with language memory may also process musical sounds.

My own behavioural research has focused on looking at the similarities between verbal and musical short-term memory using predictions from Baddeley and Hitch's working memory model (Williamson et al., in press). Although I have uncovered a number of differences in the features of short-term memory for speech and music I have noted a number of interesting similarities. For example,

I have found evidence that getting people to move their articulators (whispering) can disrupt verbal and musical recall. Previously the use of articulatory suppression was only thought to impact upon language memory, probably because it occupies the speech-motor planning system of the brain that we use to rehearse speech sounds. One interpretation of the result is that this system of mental rehearsal is actually involved in maintaining aspects of both speech and music sounds in memory.

The benefits of a musical background

One prediction that we might make if there were an overlap between language and music processing in the brain would be that expertise in one domain would benefit processing in another. But do trained musicians show any advantage in processing language?

This question is the subject of lively debate in the literature. Some of the supportive evidence has suggested that musicians may show a number of benefits to their language perception and processing. Studies have shown musicians to have small increases in general IQ score (Schellenberg, 2004), improved verbal memory (Ho et al., 2003), improved pronunciation of foreign languages (Tanaka & Nakamura, 2004), and more robust encoding of the pitch of speech