Cross-domain processing of musical and vocal emotions in cochlear implant users

Alexandre Lehmann1, 2, 3* and Sébastien Paquette2, 3

1 Department of Otolaryngology Head and Neck Surgery, McGill University, Montreal, QC, Canada, 2 International Laboratory for Brain, Music and Sound Research, Center for Research on Brain, Language and Music, Montreal, QC, Canada, 3 Department of Psychology, University of Montreal, Montreal, QC, Canada

Keywords: cross-domain processing, emotion, music, voice, cochlear implant, brain plasticity, neural overlap

Music and voice bear many similarities and share neural resources to some extent. Experience dependent plasticity provides a window into the neural overlap between these two domains. Here, we suggest that research on auditory deprived individuals whose hearing has been bionically restored offers a unique insight into the functional and structural overlap between music and voice. Studying how basic emotions (happiness, sadness, and fear) are perceived in auditory stimuli constitutes a favorable terrain for such an endeavor. We outline a possible neuro-behavioral approach to study the effect of plasticity on cross-domain processing of musical and vocal emotions, using cochlear implant users as a model of reversible sensory deprivation and comparing them to normal-hearing individuals. We discuss the implications of such developments on the current understanding of cross-domain neural overlap.

Cross-domain Neural Overlap and Plasticity

Our musical and vocal perception abilities have such a close relationship that some authors suggested that the former originated from the latter or vice-versa (Honing et al., 2015; Peretz et al., 2015). To what extent do music and voice share functional and structural networks and at which stage of auditory processing they are differentiated are open questions. Functional magnetic resonance imaging (fMRI) studies show the co-activation of brain regions with possibly distinct underlying neural populations (Peretz et al., 2015). Research on expert populations has suggested reciprocal interactions between neural circuits associated with the domains of music and voice (Patel, 2011; White-Schwoch et al., 2013; summarized by Paquette and Mignault Goulet, 2014). Indeed studies have shown that musicians have enhanced speech processing capacity, which is reflected in both cortical and subcortical neural measures (Bidelman et al., 2011, 2014; Parbery-Clark et al., 2012). Musicians can be used as a model of learning-induced plasticity to investigate how such cross-domain transfer effects unfold over time (Strait and Kraus, 2014; Strait et al., 2014). Here we argue that, sensory deprivation offers a complementary model to shed light on the plastic reorganization of brain networks involved in particular functions.

Temporary Deafened Individuals offer a Unique Insight into Auditory Neural Plasticity

Cochlear implants (CI) are bionic devices that can restore the sense of hearing in profoundly deaf individuals. We argue that cochlear implant users offer a promising model to study the mechanisms of cross-domain plasticity because they undergo different trajectories...
of auditory development: deafness of various origins results in a variable period of auditory deprivation followed by surgical restoration of auditory input and an intense rehabilitation period, yielding variable individual auditory outcomes.

Signal transmitted from the implant to the auditory nerve is impoverished compared to natural hearing. Critically, the access to pitch cues is impaired, reduced to a small number of frequency bands. As a result, cochlear implant users can potentially perceive speech relatively well in a quiet setting, but understanding it in noise, or accurately perceiving music is very challenging since both tasks rely on pitch information (Gfeller et al., 2007). Perception is not only affected by the impoverished auditory input, but also by neural re-organization following auditory deprivation, from the periphery to the cortex. In absence of auditory input, auditory nerve fibers start to degenerate and the auditory cortex can be recruited by visual and somatosensory systems (Collignon et al., 2011; Lazzouni and Lepore, 2014). Such plastic changes can prevent the auditory cortex from fully recovering its initial function after the auditory input is restored via an implant (Lee et al., 2001; Bavelier and Hirshorn, 2010; Sandmann et al., 2012; Sharma et al., 2015).

To date, little is known of the neural correlates of music and voice processing in cochlear implants and the extent to which those processes overlap. Only one study has performed a direct comparison of the neural correlates of speech and music perception in CI users. Using positron emission tomography (PET), Limb et al. (2009), reported increased activation and greater cortical recruitment in implant recipients compared to normal hearing controls, during both speech and music listening. This effect was stronger for speech—for which CI users are more proficient than music—and suggest a link between auditory performance and degree of auditory cortical activation.

**Emotion as a Cross-domain Terrain of Choice to Study Neural Overlap**

An important part of our social interaction relies on accurate emotion perception. In normal-hearing individuals, evidence from neuropsychology suggest the existence of an auditory emotional neural pathway, distinct from auditory perception, that might be shared across musical and vocal domains and have both cortical and subcortical components (Peretz, 2011). A systematic comparison of the vocal and musical domains suggests a close acoustical relationship for emotional expression, with similar emotion-specific acoustic cues patterns (Juslin and Laukka, 2003). Several of those patterns relate to the pitch dimension, such as prosody for voice (variations in the pitch contour) and melody for music. The perception of pitch is severely degraded in cochlear implant users, thus limiting their access to those important cues, but other non-pitch based cues can also convey emotions (Gabriëls and Lindström, 2010). It was recently demonstrated in amusics (individuals with a lifelong pitch perception deficit; Peretz, 2013) that non-pitch based cues (e.g., tempo, pulse clarity) can be used to identify musical emotions (Gosselin et al., 2015). These cues are available to some extent to CI users (Kong et al., 2004; Looi et al., 2012), and should allow them a certain degree of emotional perception. CI users have a documented deficit in both vocal and musical emotion recognition; emotional categories and dimensions are not uniformly impaired. They can recognize some categories of emotion in voice or music above chance, but not as well as normal hearing controls (Hopyan et al., 2012; Nakata et al., 2012; Volkova et al., 2013; Wang et al., 2013). They have difficulty perceiving arousal of musical excerpts but not valence (Ambert-Dahan et al., 2015). These differences could be due to the relatively spared abilities of CI users to perceive temporal variations, while having an impaired pitch perception. They could also reflect differences in the complexity of stimuli employed and how they are handled by speech-optimized processors, suggesting that ad-hoc stimuli are required to accurately compare the two domains. This could explain why no study has yet directly compared emotion processing in CI users across the domains of music and voice. To date, there is very little neuro-imaging evidence building up on the aforementioned behavioral findings. Only one study evaluated the impact of two implant processing strategies on the perception of prosody (Agrawal et al., 2013) and demonstrated that electroencephalography (EEG) is a useful tool to reveal differences between strategies coding specific features.

**Toward a Study of Cross-domain Processing of Musical and Vocal Emotions in Cochlear Implant Users**

A large part of the research on auditory affective processing has been conducted on prosody utilizing words or sentences spoken with various emotional expressions and complex musical pieces expressing varying degrees of emotion. It is not possible to directly compare those results between music and voice because of many confounding variables; factors such as speech semantics, length, harmony, and context are likely to recruit different neural networks. We argue that a necessary first step to study cross-domain processing of musical and vocal emotions is to use an experimental paradigm that moves away from the fairly complex sounds used in the existing literature, using stimuli that enable a controlled comparison between the domains of music and voice. A possible approach would be to use the most primitive affect expressions (primal interjections close to those of babies and animals) in each domain: non-speech vocalizations and brief mono-instrumental musical excerpts.

In the vocal domain, non-speech vocalizations (e.g., screams, laughter) depicting basic emotions that are minimally conventionalized, relatively universal and fundamental to spontaneous human communication (Scherer, 1986), could be used. Stimuli like the Montreal Affective Voices (Belin et al., 2008), consisting of short vocal interjections on the vowel /a/ expressing basic emotions, represent the most primitive form of emotion in their domain. They have minimal semantic information and minimal interaction with linguistic processes (Bestelmeyer et al., 2010). Compared to speech prosody, vocalizations are treated preferentially in the brain (Pell et al., 2015). When it comes to music, finding the most basic emotions...
and avoiding interaction with other processes require stepping away from conventional structure (limited by mode or tempo), reducing the length of the stimuli and reducing its emotional complexity. Stimuli like the Musical Emotional Bursts (Paquette et al., 2013) could be used for comparison, they consist of a few spontaneous notes on a clarinet or violin expressing basic musical emotions, they are minimally conventionalized and represent the most primitive form of emotion in their domain. They are all the more similar to vocal stimuli because they use continuous pitch instrument (e.g., the violin which offers a seamless progression between notes, giving the stimuli a quasi-vocal quality), whereas most studies have used discrete pitch instruments (e.g., the piano where one key corresponds to one pitch), which further hinders the direct comparison with vocal stimuli.

These highly similar vocal and musical stimuli seem well-suited to study cross-domain overlap in any population and their primitive quality could be extremely useful to study plasticity in CI users.

A second step would be to pair a well-controlled behavioral paradigm using those stimuli (allowing a direct comparison of musical and vocal domains) with a neuro-imaging modality that is acceptable for use with cochlear implants. Except for a few recent exceptions, implants are not MR-compatible. High-density EEG (Gilley et al., 2010; Zhang et al., 2011; Timm et al., 2014) and PET-scan (Okazawa et al., 1996; Limb et al., 2009; Lazard et al., 2010) have both been used successfully in cochlear implant users. Both methods have drawbacks: EEG recordings are contaminated by massive electrical artifacts from the implant and PET requires the injection of a radioactive isotope. Emerging as a promising brain-imaging modality for CI research is functional near-infrared spectroscopy (fNIRS). fNIRS has been successfully used to study the response to auditory stimuli in cochlear implant users (Sevy et al., 2010) and emotion-related activation in the general population (Herrmann et al., 2003; Plichta et al., 2011). This non-invasive technique measures blood oxygenation level differences using infrared light and is therefore unaffected by electrical artifacts. It is portable and has a better temporal resolution than functional MRI (Villringer and Chance, 1997). Conversely it has a worse spatial resolution and cannot access subcortical sources such as the limbic system (Kochel et al., 2011).

The proposed neuro-behavioral approach would be well-suited to study the effect of plasticity on cross-domain processing of musical and vocal emotions, using cochlear implant users as a model of reversible sensory deprivation and comparing them to normal-hearing individuals. The effect of multiple regressors could be assessed by recruiting an heterogeneous cohort of individuals spanning the continuum of factors known to affect plasticity such as the duration of auditory deprivation or the age at implantation (Lazard et al., 2012). This would represent a stepping-stone to ask further questions of interest regarding the effect of plasticity on cross-domain neural overlap. From a basic science perspective, the rationale is to understand a complex system by reverse-engineering its dysfunctions. What are the structural and functional overlaps between music and voice processing after implantation? Would the reduction of auditory cortical resources, together with the fact that music and vocal signals are more similar after being processed by the device, favor an increased neural overlap between domains? Conversely, would any remaining overlap break-down in favor of a more segregated re-organization guided by the non-pitch based, domain relevant cues?

Characterizing those mechanisms can inform novel clinical approaches, possibly through individualized rehabilitation and brain stimulation. For instance, if good performers (CI users with good speech scores) make use of overlapping structures in an optimal fashion compared to poor performers, can we boost residual neural processes in the latter group? It has been suggested that musical training can improve speech outcomes in this population (Patel, 2014), but what stages of the auditory pathway are best candidates for a cross-domain shaping of function and/or structure? Auditory features found to maximize activity of brain networks processing musical and vocal emotions in CI users could be made more salient in device processors.

Cross-domain research on cochlear implant users not only offers a unique insight into auditory neural plasticity, but also has practical implications for patients’ rehabilitation, implant design, and programming. We believe that highly comparable stimuli are needed to carry out such studies, together with an optimal imaging technique within a paradigm fine enough to reveal subtle behavioral and neural differences. Such scientific undertaking can further our understanding of how our brain processes vocal and musical emotions and how such cross-domain processing is affected by plasticity. Furthermore, such studies could provide objective measures to support the use of music in the rehabilitation of various disorders.

**Acknowledgments**

This work was supported by a CRBLM research incubator award (funded by the Fonds de Recherche Nature et Technologies and Société et Culture) to AL and by a graduate scholarship from the Canadian Institutes of Health Research to SP.

**References**

Agrawal, D., Thorne, J. D., Viola, F. C., Timm, L., Debener, S., Buchner, A., et al. (2013). Electrophysiological responses to emotional prosody perception in cochlear implant users. *NeuroImage Clin.* 2, 229–238. doi: 10.1016/j.nicl.2013.01.001

Ambert-Dahan, E., Giraud, A. L., Sterkers, O., and Samson, S. (2015). Judgment of musical emotions after cochlear implantation in adults with progressive deafness. *Name Front. Psychol.* 6,181. doi: 10.3389/fpsyg.2015.00181

Bavelier, D., and Hirshorn, E. A. (2010). I see where you’re hearing: how cross-modal plasticity may exploit homologous brain structures. *Nat. Neurosci.* 13:1309. doi: 10.1038/nn1110-1309
Belin, P., Fillion-Bilodeau, S., and Gosselin, F. (2008). The montreal affective voices: a validated set of nonverbal affect bursts for research on auditory affective processing. Behav. Res. Methods 40, 531–539. doi: 10.3758/BRM.40.2.231

Bestelmeyer, P. E., Rouger, J., DeBruine, L. M., and Belin, P. (2010). Auditory adaptation in vocal affect perception. Cognition 117, 217–223. doi: 10.1016/j.cognition.2010.08.008

Bidelman, G. M., Gandour, J. T., and Krishnan, A. (2011). Cross-domain effects of music and language experience on the representation of pitch in the human auditory brainstem. J. Cogn. Neurosci. 23, 425–434. doi: 10.1162/jocn.2009.21362

Bidelman, G. M., Weiss, M. W., Moreno, S., and Alain, C. (2014). Coordinated plasticity in brainstem and auditory cortex contributes to enhanced categorical speech perception in musicians. Eur. J. Neurosci. 40, 2662–2673. doi: 10.1111/ene.12627

Collignon, O., Vandewalle, G., Voss, P., Alibouy, G., Charbonneau, G., Lassonde, M., et al. (2011). Functional specialization for auditory-spatial processing in the occipital cortex of congenitally blind humans. Proc. Natl. Acad. Sci. U.S.A. 108, 4435–4440. doi: 10.1073/pnas.1013928108

Gabrieliess, A., and Lindström, E. (2010). The role of structure in the musical expression of emotions. Handb. Music Emot. Theory Res. Appl. 367–400. doi: 10.1093/acprof:oso/9780199230143.003.0014

Gfeller, K., Turner, C., Olsen, J., Zhang, X., Gantz, B., Froman, R., et al. (2007). Accuracy of cochlear implant recipients on pitch perception, recognition, and speech reception in noise: Ear Hear. 28, 412–423. doi: 10.1097/AUD.0b013e3180479318

Gilley, P. M., Sharma, A., Mitchell, T. V., and Dormann, M. F. (2010). The influence of a sensitive period for auditory-visual integration in children with cochlear implants. Restor. Neurol. Neurosci. 28, 207–218. doi: 10.3233/RNN-2010-0525

Gosselin, N., Paquette, S., and Peretz, I. (2015). Sensitivity to musical emotions in congenital amusia. Cortex 71, 171–182. doi: 10.1016/j.cortex.2015.06.022

Herrmann, M. J., Ehls, A.-C., and Fallgatter, A. J. (2003). Prefrontal activation through task requirements of emotional induction measured with NIRS. Biol. Psychol. 64, 255–263. doi: 10.1016/S0301-0511(03)00095-4

Honig, H., Cate, C., ten, Peretz, I., and Trehub, S. E. (2015). Without it no music: cognition, biology and evolution of musicality. Philos. Trans. R. Soc. Lond. B Biol. Sci. 370:20140088. doi: 10.1098/rstb.2014.0088

Hoppyan, T., Peretz, I., Chan, L. P., Papsin, B. C., and Gordon, K. A. (2012). Children using cochlear implants capitalize on acoustical hearing for music perception. Front. Hum. Neurosci. 3:425. doi: 10.3389/fnhum.2012.00425

Justlin, P. N., and Laukka, P. (2003). Communication of emotions in vocal expression and music performance: different channels, same code? Psychol. Bull. 129, 770–814. doi: 10.1033/2099-1295.7.770

Kochel, A., Plichta, M. M., Schäfer, A., Leutgeb, V., Scharmüller, W., Fallgatter, A. J., et al. (2011). Affective perception and imagery: a NIRS study Int. J. Psychophysiol. 80, 192–197. doi: 10.1016/j.ijspsycho.2011.03.006

Kong, Y.-Y., Cruz, R., Jones, J. A., and Zeng, F.-G. (2004). Music perception with speech? The OPERA hypothesis. Aud. Cogn. Neurosci. 2:142. doi: 10.3389/fpsyg.2011.00142

Patel, A. D. (2011). Why would musical training benefit the neural encoding of speech? The OPERA hypothesis. Hear. Res. 308, 98–108. doi: 10.1016/j.heares.2013.08.011

Pell, M. D., Rothermich, K., Liu, P., Paulmann, S., Sethi, S., and Rigoulot, S. (2015). Preferential decoding of emotion from human non-linguistic vocalizations versus speech prosody. Biol. Psychol. 111, 14–25. doi: 10.1016/j.biopsycho.2015.08.008

Peretz, I., Vuvan, D., Lagros, M.-E., and Armony, J. L. (2015). Neural overlap in processing music and speech. Philos. Trans. R. Soc. B Biol. Sci. 370, 20140090–20140090. doi: 10.1098/rstb.2014.0090

Peretz, I. (2011). “Towards a neurobiology of musical emotions,” in Handbook of Music and Emotion: Theory, Research, Applications, eds P. N. Juslin and J. A. Sloboda (Oxford: Oxford University Press), 99–126.

Patel, A. D. (2013). “The biological foundations of music: insights from congenital Amusia,” in The Psychology of Music 3rd Edn. ed D. Deutsch (Academic Press), 551–564.

Plichta, M. M., Gerdes, A. B. M., Alpers, G. W., Harnisch, W., Brill, S., Wieser, M. J., et al. (2011). Auditory cortex activation is modulated by emotion: a functional near-infrared spectroscopy (NIRS) study. Neuroimage 55, 1200–1207. doi: 10.1016/j.neuroimage.2011.01.011

Sandmann, P., Diller, N., Eichele, T., Meyer, M., Kegel, A., Pascual-Marqui, R. D., et al. (2012). Visual activation of auditory cortex reflects maladaptive plasticity in cochlear implant users. Brain J. Neurol. 135, 555–568. doi: 10.1016/j.brainimg.2012.05.022

Scherer, K. R. (1986). Vocal affect expression: a review and a model for future research. Psychol. Bull. 99, 143–165. doi: 10.1037/0033-2909.99.2.143

Sevy, A. B. G., Bortfeld, H., Huppert, T. J., Beauchamp, M. S., Tonini, R. E., and Oghalai, J. S. (2010). Neuroimaging with near-infrared spectroscopy demonstrates speech-evoked activity in the auditory cortex of deaf children following cochlear implantation. Hear. Res. 270, 39–47. doi: 10.1016/j.heares.2010.09.010

Sharma, A., Campbell, J., and Cardon, G. (2015). Developmental and cross-modal plasticity in deafness: evidence from the P1 and N1 event related potentials in cochlear implanted children. Int. J. Psychophysiol. 95, 135–144. doi: 10.1016/j.ijspsycho.2014.04.007

Strait, D. L., and Kraus, N. (2014). Biological impact of auditory expertise across the life span: musicians as a model of auditory learning. Hear. Res. 308, 109–121. doi: 10.1016/j.heares.2013.08.004

Strait, D. L., O’Connell, S., Parbery-Clark, A., and Kraus, N. (2014). Musicians’ enhanced neural differentiation of speech sounds arises early in life: developmental evidence from ages 3 to 30. Cereb. Cortex 24, 2512–2521. doi: 10.1093/cercor/bht103

Timm, L., Vuust, P., Brattico, E., Agrawal, D., Dehener, S., Buechner, A., et al. (2014). Preferential decoding of emotion from human non-linguistic vocalizations: evidence from the P1 and N1 event related potentials in cochlear implanted children. Int. J. Psychophysiol. 81, 127–137. doi: 10.1016/j.ijspsycho.2014.01.018

Villringer, A., and Chance, B. (1997). Non-invasive optical spectroscopy and imaging of human brain function. Trends Neurosci. 20, 435–442. doi: 10.1016/S0166-2236(97)01132-6
Volkova, A., Trehub, S. E., Schellenberg, E. G., Papsin, B. C., and Gordon, K. A. (2013). Children with bilateral cochlear implants identify emotion in speech and music. Cochlear Implants Int. 14, 80–91. doi: 10.1179/1754762812Y.0000000004

Wang, D. J., Trehub, S. E., Volkova, A., and van Lieshout, P. (2013). Child implant users' imitation of happy- and sad-sounding speech. Front. Psychol. 4:351. doi: 10.3389/fpsyg.2013.00351

White-Schwoch, T., Carr, K. W., Anderson, S., Strait, D. L., and Kraus, N. (2013). Older adults benefit from music training early in life: biological evidence for long-term training-driven plasticity. J. Neurosci. 33, 17667–17674. doi: 10.1523/JNEUROSCI.2560-13.2013

Zhang, F., Hammer, T., Banks, H.-L., Benson, C., Xiang, J., and Fu, Q.-J. (2011). Mismatch negativity and adaptation measures of the late auditory evoked potential in cochlear implant users. Hear. Res. 275, 17–29. doi: 10.1016/j.heares.2010.11.007

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2015 Lehmann and Paquette. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.