PITCH PROCESSING IN MUSIC AND SPEECH

Barbara Tillmann[1,2,3]

[1] **CNRS, UMR5292; INSERM, U1028; Lyon Neuroscience Research Center, Auditory Cognition and Psychoacoustics team, Lyon, F-69000, France**

[2] **University of Lyon 1, Lyon, F-69000, France**

[3] **MARCS Institute, University of Western Sydney**

barbara.tillmann@cnrs.fr

The present paper proposes an overview of research that investigates pitch processing by considering cognitive processes (related to context, learning, memory and/or knowledge) for both music and language materials. Research investigating cross-domain influences of expertise (either in music or tone languages) and deficits (as in congenital amusia), referred to as positive and negative transfer effects, also contributes to our understanding of domain-specificity or –generality of mechanisms involved in pitch processing.

INTRODUCTION

A highly-debated question is to what extent music and language share processing components. Beyond syntax and temporal structure processing, one studied aspect is pitch-processing in a given domain and across domains (e.g., [1]).

Pitch processing is crucial in music. For example, in Western tonal music, it is a form-bearing dimension (next to temporal structures). Pitch processing is also crucial in speech, notably for discriminating questions and statements, as well as for communicating emotional expressions. This is valid for non-tone intonation languages (e.g., English, French) and tone languages (e.g., Mandarin, Thai, Vietnamese). However, for tone languages, pitch processing is even more crucial as pitch information is used for communicating word meaning. Tone languages comprise 70% of the world's languages and are spoken by more than 50% of the world's population. In these languages, tone variations (comprising predominantly fundamental frequency (F0) height and contour parameters) at the syllabic level have the same effect on word meaning as do vowel and consonant variations in non-tone languages. For example, the syllable /ma/ combined with different tones (e.g., tones describing contours of rather constant level, rising dipping or falling patterns in Mandarin) represents different lexical items.

Interestingly, music and language differ in the size of the pitch differences that are relevant for each of the systems (speech intonation versus musical structures). For speech intonation of non-tonal languages, F0 variations, in particular those indicating statements and questions, are typically coarse (up to more than 12 semitones[1] for the pitch rise of the final word in a question; e.g., [2]). For music (as in most research, we are referring here mostly to the Western tonal system), however, the pitch variations are typically more fine-grained (1 or 2 semitones; see [3]). In tone languages, the range of F0 variations can be as small as in music of the Western tonal system or larger depending on the tones and the tone languages.

The present paper proposes an overview of research that investigates pitch processing by considering cognitive processes (related to context, learning, memory and/or knowledge) for both music and language material. While extensive research has focussed on pitch processing in musical material as well as the influence of musical expertise (e.g., comparing musicians and nonmusicians), research investigating pitch processing in tone language material for either musicians, nonmusicians or tone language speakers (or both), provides complementary information about underlying mechanisms. Furthermore, investigating deficits (such as in amusia) provides further insights into the potential domain-specificity or domain-generality of pitch processing mechanisms.

PITCH PROCESSING: INFLUENCE OF CONTEXT, KNOWLEDGE AND MATERIAL

Pitch processing has been investigated in detail within the research domain of psychoacoustics. However, psychoacoustic studies investigating early processes of pitch discrimination have mostly used pure tones or complex tones (see [4] for a review) and the rare studies using verbal material focused on vowel formants or one of the formants (e.g., [5]).

Even though no study has investigated pitch discrimination thresholds within the same participants for both non-verbal and verbal materials, some studies have examined thresholds for each domain separately. The findings suggest that pitch discrimination is more precise for musical material than for speech material (in typical listeners). For example, pitch discrimination thresholds (above 100Hz) are around 0.2% for complex tones [4] and are 10 times larger (2%) for vowels [6]. A recent study has compared for the first time the same series of pitch changes for music and speech in the same listeners

FOOTNOTE

[1] A semitone (or half tone) is the smallest musical interval commonly used in Western tonal music (e.g., the interval between tone C and tone C#); one semitone corresponds to 100 cents.

(even though it was not measuring thresholds). The findings of typical nonmusician listeners are consistent with these previous data, notably reporting higher accuracy for pitch processing when instantiated in musical (non-verbal) material than in verbal material [7]. Interestingly, individuals afflicted by congenital amusia (i.e., a lifelong deficit of music processing) showed better performance for verbal material than for musical material, even though they were generally impaired (see below). This benefit of speech for pitch discrimination in congenital amusia (and the reverse data pattern for typical nonmusician listeners) might be due to differences at early perceptual processing steps (e.g., exploiting differently the energy distribution in the sounds' spectrum linked to the presence vs. absence of formants) and/or higher level processing steps (i.e., strategic influences, attention and memory) (see [7] for further discussion).

Another characteristic of numerous psychoacoustic studies is the investigation of pitch processing of a single sound, which is presented in an isolated way without other surrounding sounds, thus “without a sound context” (besides those of the experimental context of the paradigm). In contrast, research in cognitive psychology has reported powerful effects of sound contexts as well as knowledge-driven influences (also referred to as top-down influences) on perceptual processing. These cognitive influences interact with the stimulus-driven influences (also referred to as bottom-up influences). Some research strands have investigated these influences on pitch processing in particular. In a natural environment, a sound rarely occurs on its own, but it rather occurs in a context with other sounds (e.g., presented in a sequence). The context may contain local regularities (such as a repeating pattern, e.g., AABAABAAB) or may refer to a structural system, which is based on statistical regularities (i.e., how the events are used together and thus define regularities of frequencies of occurrence or co-occurrence, for example). If listeners have knowledge about the underlying sound system, this knowledge allows listeners to develop expectations for future incoming sound events, and these expectations then influence event processing. Musical systems, which include structural regularities based on events differing in pitch, and listeners’ knowledge thereof are examples of these cognitive influences on pitch processing.

Cognitive influences for the processing of musical materials have been reported not only for musician listeners, but also for nonmusician listeners who have been reported to have knowledge about the musical system of their culture (e.g., [8]). Performance of nonmusicians is compared to performance of musicians, that is individuals who are trained in performing music, with theoretical background knowledge and several thousands of hours of music production. While for some tasks, nonmusicians perform as musicians, for other tasks they are outperformed by the musicians. The musicians’ musical background has been shown to influence pitch processing not only for musical, non-verbal material, but also for speech material. Conversely, it is interesting to also investigate the influence of language background on pitch processing in speech and musical material – in particular, native speakers of tone languages. This research thus also integrates into the research domain investigating whether processes and/or cognitive and neural correlates are domain-specific or domain-general.

CONTEXT EFFECTS: TOP-DOWN, KNOWLEDGE-DRIVEN EFFECTS ON PITCH PROCESSING

The influence of context effects on pitch processing has been shown not only for musical materials, but also for non-musical materials, that is, for example, contexts made out of several tones defining Gestalt-like features (e.g., a series of tones describing a descending melodic contour). When a sound context is defined by such a descending melodic contour (i.e., sounds constantly decreasing in pitch height from high to low frequency), participants are more sensitive to the detection of a target whose pitch is placed in the continuity of the instilled contextual movement than when it violates this movement [9]. The perceptual expectations developed with contextual movement function as an indirect signal, which facilitates the detection of the target sound. Music cognition research has investigated whether expectations based on listeners’ tonal knowledge might also serve as an indirect indication of the target’s pitch height and thus facilitate its processing. More specifically, processing the pitch of a tonally strongly related tone at the end of the melody (and thus supposed to be expected tone) should be facilitated in comparison to processing the pitch of an unexpected or less-related tone.

For musician listeners, Francès [10] has shown that tonal expectations influence listeners’ perceptual ability to detect changes in the shift of the F0 of a musical tone (the target), which was followed by another tone (thus presented as a tone pair, defining an interval). When presented without additional, surrounding tones, the performed mistunings of the target tone, which either reduced or increased the pitch interval defined by the two tones, were detected by the participants (all musicians). However, when placed in a musical context (i.e., additional tones, which defined tonal structures and a tonality, were presented before the same tone pair), the same mistunings of the target tone were only perceived when they increased the pitch intervals and conflicted with listeners’ expectations linked to musical anchoring, but not when they were in agreement with the musical patterns of tension and relaxation induced by the musical structures (see also [11]).

For nonmusician listeners, Warrier and Zatorre [12] have shown that an increasingly long tonal context (i.e., melodies with increasing duration) improved participants’ capacity to process pitch information despite timbre changes. The authors interpreted this benefit by suggesting that the contextual tones create a stronger reference point from which to judge the pitch, and that the tonal structure of the melodies provides cues that facilitate pitch processing. In the experimental material, the to-be-judged tone played the role of the tonic, the most referential tonal function, and this function might have been also beneficial for pitch processing.

Other research focuses more specifically on the influence of tonal functions and structures on pitch processing, and this in particular for nonmusician listeners (see [8, 13] for reviews). Most of this research investigated harmonic structures by using chord sequences, but some more recent research applied these research questions to melodic materials and to tonal structures implemented in melodies, thus allowing for the investigation of pitch processing more specifically. In Marmel et al. [14], the influence of tonal function (also referred to as tonal relatedness) was investigated for pitch processing in melodies. Tonal expectations for target tones were manipulated in melodic contexts by changing just one tone of the context (which could be repeated). This subtle manipulation changed the tonal function of the final tone. It allows investigation of the influence of tonal expectations on pitch perception while controlling expectations based on information stored in sensory memory buffer. Excluding this kind of sensory influence (which otherwise would be a parsimonious, stimulus-driven explanation) allows for the investigation of cognitive influences, thus to provide evidence for listeners' tonal knowledge, which influences perception. Results showed that even for nonmusician listeners, the tonal relatedness of a target tone influences not only listeners' subjective judgments of tuning/mistuning (using a subjective rating scale), but also the speed of processing: in-tune tones that are tonally related are processed faster than in-tune tones that are less related (shown by using a priming task). Most interestingly, the tonal expectations also influence pitch discrimination: participants' performance in a pitch comparison task requiring the processing of small mistunings was better when the to-be-compared tones were tonally related to the melodic context (i.e., functioning as a tonic tone rather than as a subdominant tone). The findings suggest facilitation of early perceptual processing steps via knowledge- and attention-related processes (and not only later cognitive processing steps related to, for example, decision making for the experimental task). This has been further confirmed with finer differences in tonal relatedness (excluding the central tonic; [15]) and with even more controlled experimental material, using melodies played by pure tones to strip off potential sensory influences of the complex harmonic spectrum of the musical timbre used to play the context [16]. It is worth underlying that these findings have been obtained for nonmusician listeners. These tonal context effects do not require explicit musical knowledge, but point to the power of implicit cognition (here listeners' knowledge about the musical system, just acquired by mere exposure; see [8, 17]). The early influence of tonal expectations has been further supported by results of an Evoked-Related Potential study: tonal expectations modulated tone processing within the first 100 msec after tone onset, resulting in an Nb/P1 complex that differed in amplitude between tonally related and less-related conditions. The results suggest that cognitive tonal expectations can influence pitch perception at several steps of processing, starting with early attentional selection of pitch [18].

Benefits of tonal knowledge on pitch processing have also been shown for pitch structure knowledge newly acquired in the laboratory. A recent behavioural study combined implicit learning and priming paradigms [19]. Participants were first exposed to structured tone sequences without being told about the underlying artificial grammar of the sequences. They then made speeded judgements on a perceptual feature of target tones in new sequences (i.e., in-tune/out-of-tune judgements). The target tones respected or violated the structure of the artificial grammar and were thus supposed to be expected or unexpected in that grammatical framework. Results of this priming task revealed that grammatical tones were processed more quickly and more accurately than ungrammatical ones. These findings show that top-down expectations based on newly acquired structure knowledge (i.e., acquired in the lab) influence processing speed (i.e., response times) of the target tones. It remains to be shown whether these top-down expectations can go beyond this influence and are powerful enough to facilitate early perceptual processing steps (e.g., pitch processing per se) and not only processes linked to decisions and other task-related processes.

Tonal structures and listeners' knowledge thereof does not influence only performance in perceptual tasks focusing on pitch, but also performance in memory tasks, which require processing of the pitch dimension (e.g., comparing two tones or two tone sequences separated by a delay and indicating whether these are the same or different). Participants show better memory performance for tonal compared to atonal chord sequences and melodies [20-21]. Tonal sequences are tone sequences that respect Western musical regularities; atonal sequences are those that do not. The benefit of tonal structures on memorizing tone sequences has been shown for both Western nonmusicians and musicians [22]. The benefit was observed when the task required maintenance of tone information, but not when manipulation was required (comparing two sequences and judging whether they were same or different, with "same" being defined as all tones played correctly in the backward order).

However, for the simpler task (requiring only maintenance, that is by comparing both sequences with tones in the same (forward) order), the benefit of the tonal structure on short-term memory was reduced for individuals with congenital amusia who have been reported to have deficits in music perception and production. Congenital amusia is a life-long deficit of music processing without brain damage or sensory deficits. Individuals with congenital amusia have difficulties recognizing familiar tunes without lyrics and detecting an out-of-key or out-of-tune note. This musical disorder occurs despite normal performance on tests of intelligence, auditory processing, cognitive functioning, language processing, and it is not due to a lack of environmental stimulation to music (see [23-25] for extensive testing).

This condition has been described as being based on impaired processing of the pitch dimension, notably with a deficit of pitch perception (e.g., [24]), but particularly of pitch memory (e.g., [26]). When tested for short-term memory with tone sequences containing tonal structure, this population did not show the benefit of tonal sequences over atonal sequences in terms of accuracy. However a benefit was shown for response times, notably with faster response times for tonal sequences than for atonal sequences (as observed for the control participants). These findings suggest that some implicit processing of tonal structures is potentially preserved in congenital amusia, which can also influence pitch memory [27]. This observation conforms with data of other studies suggesting implicit processing of pitch despite congenital amusia (e.g., [28-29]).

This section reviewed findings for top-down influences due to listeners' tonal knowledge or newly learned tone structure knowledge based on an artificial grammar. Future research now needs to further investigate the kind of top-down influences that are driven by listeners' knowledge of linguistic structure from the language of their culture. Some research has investigated this for tone languages where the pitch dimension is crucial, with pitch carrying meaning (see next section), but no research has investigated the influence of knowledge based on context and/or whether question or statement, which can be indicated by pitch markers, will be presented.

DOMAIN SPECIFICITY OF PITCH PROCESSING IN MUSICAL AND VERBAL MATERIAL?

Regarding the debate of domain-specific or domain-general processing, some findings have suggested common pitch processing mechanisms in music and speech, notably by showing some beneficial influences (positive transfer) across domains (i.e., music and speech) due to expertise in music or in tone languages.

Musical training or expertise has been shown to improve pitch perception not only in musical contexts, but also in speech contexts. For example, musicians show improved pitch processing for prosody of non-tonal language material [30-31] and for tone-language material, such as Thai tones [32-33] and Mandarin tones [34-38]. Comparing musicians and nonmusicians is informative, but also raises the criticism that differences between the two populations might not be due to musical training, but have rather existed before starting to learn music. Longitudinal studies of musical training in the short term, e.g., within the experimental framework, have started to investigate this issue to reject the raised criticism: Nonmusician children are allocated to two groups, for example musical training versus some other kind of training (painting, drama) for several months. Comparing the performance of the children before and after training as well as between the groups after training allows investigation of the potential effects of musical training on neural correlates (anatomical, functional) and sensory and cognitive processes involved in music processing as well as in language processing. For example, Moreno et al. [39] reported that after musical training, the children of the music group processed better small pitch changes (for music and speech materials) than did the children of the painting group. And this benefit for pitch processing in speech was observed not only for their mother tongue, but also for a different, foreign language – a phenomenon that could facilitate the learning of new languages [40].

These studies have all focused on the Western tonal system and compared Western tonal musicians to nonmusicians. This thus reveals another research area where cultural investigation bias needs to be overcome (see [41] for a discussion). Notably, it would be interesting to test whether musicians who are experts for musical systems that are based on microtonal structures (that is, the octave is divided in more than 12 semitones, thus containing smaller intervals than a semitone, as for example in traditional Indian music or some African musical systems) would be even better in pitch processing for both musical and verbal materials.

Similarly to musical expertise, expertise or training in a tonal language can facilitate pitch perception and production with musical materials: Mandarin, Vietnamese and Cantonese speakers have been found to be more accurate at imitating musical pitch and discriminating intervals than English speakers ([42], see also [43]), as can be also reflected in subcortical pitch tracking (e.g., [44]). The influence of tone language background has been mostly observed for relative pitch processing (e.g., intervals, contours). Stevens et al. [45] more specifically investigated pitch contour processing in spoken Thai and English items (speech task) as well as in matched musical items for participants with tonal (Thai) and non-tonal (Australian English) language backgrounds. The overall findings suggest that expertise in tonal language leads to perceptual attunement for contour processing in spoken items as well as in musical items (even though here restricted to the speed of processing rather than extending to accuracy of processing). However, the influence of tone language background might also lead to difficulties in pitch contour processing when non-speech target sounds resemble features of linguistic tones [46]. In contrast to these results for relative pitch processing, it has been shown that listeners with tone language background did not differ from listeners with non-tone language background for absolute pitch discrimination of non-speech sounds (e.g., [42, 46]). Interestingly, in musically-trained participants, there is a link between tone language background and single pitch processing: absolute pitch (i.e., the ability to label a tone without a reference pitch) appears to be more prevalent among tone language speakers than among non-tone language speakers [47-48].

Regarding potential neural correlates of these expertise effects and their cross-domain effects, it has been proposed that musical training might shape basic sensory circuitry as well as corticofugal tuning of the afferent system, which is context-general and thus also has positive side-effects on linguistic pitch processing (e.g., [38]). Similar findings suggesting experience-dependent corticofugal tuning have been recently reported for the effects of tone language expertise on musical pitch processing [35].

In contrast to these expertise/training-related improvements of pitch processing from one domain to the other, recent research has investigated the influence of a pitch perception deficit, which has been first described for music (as in the condition of congenital amusia), on pitch perception in speech. This could be also labelled as "negative transfer" – in parallel to the positive transfer and benefit of expertise, discussed above.

As introduced above, congenital amusia has been thought to result from a musical pitch-processing disorder. Individuals with congenital amusia have impaired perception of pitch directions for pure tones [25] and for detecting pitch deviations that are smaller than two semitones in sequences of piano notes [49] as well as in note pairs [24]. Initial reports have suggested that the deficit was restricted to pitch processing in music, and did not extend to pitch processing in speech material. Individuals with congenital amusia have been reported to be unimpaired in language and prosody tasks, such as learning and recognizing lyrics, classifying a spoken sentence as statement or question based on final falling or rising pitch information (e.g., [23-24]).

However, more recent studies revealed deficits also for pitch processing in amusia – for prosody and for tone languages. Amusics showed mild deficits in processing speech intonation (questions vs. statement) or emotional prosody in their native language (English or French; [2, 50-51]), or in processing pitch contrasts in tone language words (Mandarin or Thai) - whether for native Mandarin speakers [52-54] or native French speakers [55-56]. Note that when tested with natural speech, which involves multiple acoustic cues, Mandarin people with amusia were not impaired [57-58]. Interestingly, people with amusia who are native speakers of Cantonese show better pitch processing abilities than those with amusia who are native speakers of non-tonal languages (English, French; [38]). For non-tonal language speakers, it has been shown that people with amusia performed better with speech than with musical analogues, especially for those individuals with amusia and high pitch discrimination thresholds (over one semitone), even though they were also impaired for speech material in comparison to the control participants. This data pattern was observed both for tone language material (Thai) and a single repeated syllable (/ka/), both in comparison to their non-verbal/musical analogues [7, 56]. Nevertheless, for both verbal and musical materials, French-speaking people with amusia were impaired in comparison to their control participants. In conclusion, speech may enhance pitch processing in amusia, even though it does not necessarily restore normal processing.

In addition to these positive and negative transfer effects (due to expertise (in music or tone language) or deficit (in congenital amusia)) on pitch processing in music and speech materials, pitch processing capacities have been recently linked to phonological and phonemic awareness abilities [60-61]. Some findings have led to the hypothesis that there might be shared or common neural bases for pitch-related impairments in amusia/tone-deafness and phonemic awareness (i.e., the ability to manipulate phonemes and syllables in spoken words) in dyslexia [61].

TRAINING OF PITCH PROCESSING: CROSS-DOMAIN AND CROSS-MODAL EFFECTS

The previous section has reviewed some research investigating pitch processing in both verbal and musical materials. Expertise and training have beneficial effects for pitch processing in both music and language. And even though individuals with congenital amusia have pitch-processing deficits in both domains, the deficits are less pronounced for the verbal material (at least for the individuals with amusia with higher pitch discrimination thresholds). Based on this finding, one might thus wonder in how far it might be possible to exploit this observation to train pitch processing – that is to train pitch processing with language material, aiming for an improvement of pitch processing in musical material. However, some previous findings also suggest that pitch processing/learning is not independent in music and speech: For example, Wong and Perrachione [62] reported an association between participants' ability to learn pitch patterns in syllables, their ability to perceive pitch patterns in non-lexical contexts and their previous musical experience.

Regarding training and rehabilitation for pitch processing, another approach exploits beneficial effects of audio-visual integration. It has been previously shown that the combination of sensory information across senses can modify perception. The simultaneous presentation of an auditory signal has been shown to improve visual performance for various tasks, even when the auditory signal was not informative regarding the visual task (e.g., [63-64]). Interestingly, these benefits based on cross-modal interactions are maximally effective when the perception of one of the signals (i.e., in only one modality) is weak. Caclin et al. [64], for example, reported the benefit of an (uninformative) auditory cue on visual processing in particular for myopic participants with poor visual performance.

Another population (beyond those with amusia) with strong deficits in pitch processing are hearing-impaired patients with cochlear implants. Cochlear implants are surgically implanted devices that directly stimulate the auditory nerve in individuals with profound deafness. However, current technology is limited in transmitting spectral information, which leads to impoverished pitch processing, thus affecting both music and speech processing (i.e., prosody). Galvin et al. [65] have started to propose training programs with short musical sequences, which are presented in addition to visual cues (informative for pitch and contour). Using this kind of display combining auditory and visual information in training, has led to improved melodic contour identification tasks (also for new contours, thus showing some kind of generalization) in patients with cochlear implants.

More recently, audio-visual interactions and benefits have been exploited for testing individuals with congenital amusia. Albouy et al. [66] investigated whether audio-visual facilitation can be observed in congenital amusia, notably by presenting uninformative visual cues in addition to the to-be-processed sound sequences (requiring pitch change detection). Results revealed that individuals with amusia and control participants benefited from simultaneous visual information: accuracy was higher and response times shorter in the audiovisual condition than in the auditory-only condition. These findings suggest that individuals with amusia can benefit from multisensory integration to improve pitch processing. The results thus provide the first step towards the possibility of exploiting multisensory paradigms to help reducing pitch-related deficits in congenital amusia.

CONCLUSION

The research reviewed here investigates pitch processing for music and speech materials, with a focus on the influence of cognitive processes (related to context, learning, memory, knowledge and/or expertise). The strength of cognitive influences (based on listeners' knowledge) on pitch perception (even down to early attentional selection of pitch) has been shown in particular for musical materials. Complementary information for our understanding of pitch processing and the domain-specificity or –generality of potentially involved mechanisms has been provided by research investigating cross-domain influences of expertise (either in music or tone languages) and deficits (as in congenital amusia). Results rather point to mostly domain-general mechanisms or shared mechanisms, with some specificities in pitch processing depending on the material, which need to be further investigated. As pointed out above, most research suffers from a cultural investigation bias; this is particularly the case for music cognition research, which focuses on Western tonal music, while some research on pitch processing in speech also includes tonal languages (and not only non-tonal ones). Future research should thus open up to the investigation of pitch processing in other cultural materials and situations, as previously investigated by research on emotional connotation in music and speech (including the role of pitch) [67].

ACKNOWLEDGEMENTS

This work was supported partly by a grant from the "Agence Nationale de la Recherche" (ANR) of the French Ministry of Research ANR-11-BSH2-001-01, and conducted in the framework of the LabEx CeLyA ("Centre Lyonnais d'Acoustique", ANR-10-LABX-0060).

REFERENCES

[1] Patel, A. D. (2008). *Music, language, and the brain*. NY: Oxford University Press.

[2] Patel, A. D., Wong, M., Foxton, J., Lochy, A., & Peretz, I. (2008). Speech intonation perception deficits in musical tone deafness (congenital amusia). *Music Perception*, **25**, 357-368.

[3] Vos, P. G., and Troost, J. M. (1989). Ascending and descending melodic intervals: statistical findings and their perceptual relevance. *Music Perception*, **6**, 383-396.

[4] Moore, B. C. J. (2008). Basic auditory processes involved in the analysis of speech sounds. *Philosoph Transact Royal Soc B Biol Sci*, **363**, 947-963.

[5] Lyzenga, J., and Horst, J. W. (1995). Frequency discrimination of band-limited harmonic complexes related to vowel formants. *J. Acoust. Soc. of Am.* **98**, 1943-1955.

[6] Smith, D. R. R., Patterson, R. D., and Turner, R. (2005). The processing and perception of size information in speech sounds. *J. Acoust. Soc. Am.*, **117**, 305-318.

[7] Tillmann, B., Rusconi, E., Traube, C., Butterworth, B., Umilta, C. & Peretz, I. (2011). Fine-grained pitch processing of music and speech in congenital amusia. *Journal of the Acoustical Society of America*, **130**, 4089-4096.

[8] Bigand, E., & Poulin-Charronnat, B. (2006). Are we "experienced listeners"? A review of the musical capacities that do not depend on formal musical training. *Cognition*, **100**, 100-130.

[9] Howard, J. H., O'Toole, A. J., Parasuraman, R., & Bennett, K. B. (1984). Pattern-directed attention in uncertain-frequency detection. *Perception & Psychophysics*, **35**, 256-264

[10] Francès, R. (1958). *La perception de la musique* (2° ed.). Paris: Vrin.

[11] Bharucha, J. J. (1996). Melodic anchoring. *Music Perception*, **13**, 383-400.

[12] Warrier, C. M., & Zatorre, R. J. (2002). Influence of tonal context and timbral variation on perception of pitch. *Perception & Psychophysics*, **64**, 198-207.

[13] Tillmann B. (2005). Implicit investigations of tonal knowledge in nonmusician listeners. *Ann N Y Acad Sci.*, **1060**, 100-110.

[14] Marmel, F., Tillmann, B. & Dowling, W. J. (2008). Tonal expectations influence pitch perception. *Perception & Psychophysics*, **70**, 841-852.

[15] Marmel, F. & Tillmann, B. (2009). Tonal priming beyond tonics. *Music Perception*, **26**, 211-221.

[16] Marmel, F., Tillmann, B., Delbé, C. (2010) Priming in melody perception: tracking down the strength of cognitive expectations. *Journal of Experimental Psychology: Human Perception & Performance*, **36**, 1016-1028

[17] Tillmann, B., Bharucha, J. J., & Bigand, E. (2000). Implicit learning of music: A Self-Organizing Approach. P*sychological Review*, **107**, 885–913.

[18] Marmel, F., Perrin, F. & Tillmann, B. (2011). Tonal expectations influence early pitch processing: Evidence from ERPs. *Journal of Cognitive Neuroscience*, **10**, 3095-104.

[19] Tillmann, B. & Poulin-Charronnat, B. (2010) Auditory expectations for newly acquired structures. *Quarterly Journal of Experimental Psychology*, **63**, 1646-1664.

[20] Bharucha, J. J., & Krumhansl, C. L. (1983). The representation of harmonic structure in music: hierarchies of stability as a function of context. *Cognition*, **13**, 63-102.

[21] Dowling, W. J. (1991). Tonal strength and melody recognition after long and short delays. *Perception & Psychophysics*, **50**, 305-313.

[22] Schulze, K., Dowling, W. J., Tillmann, B. (2012). Maintenance and Manipulation of tonal and atonal sequences in non-musician and musician listeners. *Music Perception*, **29**, 255-267.

[23] Ayotte, J., Peretz, I., and Hyde, K. L. (2002). Congenital amusia: A group study of adults afflicted with a music-specific disorder. *Brain* **125**, 238-251.

[24] Peretz, I., Ayotte, J., Zatorre, R. J., Mehler, J., Ahad, P., Penhune, V. B., and Jutras, B. (2002). Congenital amusia: a disorder of fine-grained pitch discrimination. *Neuron* **33**, 185-191.

[25] Foxton, J. M., Dean, J. L., Gee, R., Peretz, I., and Griffiths, T. D. (2004). Characterization of deficits in pitch perception underlying 'tone deafness'. *Brain* **127**, 801-810.

[26] Tillmann, B., Schulze, K. & Foxton, J. (2009). Congenital amusia: A short-term memory deficit for nonverbal, but not verbal sounds. *Brain & Cognition*, **71**, 259-264.

[27] Albouy, P., Schulze, K., Caclin, A. & Tillmann, B. (2013). Does tonality boost short-term memory in congenital amusia? *Brain Research*, **1537**, 224-32

[28] Peretz I., Brattico E., Järvenpää M., and Tervaniemi, M. (2009). The amusic brain: in tune, out of key, and unaware. *Brain*, **132**, 1277-1286.

[29] Tillmann, B., Gosselin, N., Bigand, E. & Peretz, I. (2012). Priming paradigm reveals harmonic structure processing in congenital amusia. *Cortex*, **48**, 1073-1078.

[30] Magne, C., Schön, D. & Besson, M. (2006). Musician children detect pitch violations in both music and language better than nonmusician children: behavioral and electrophysiological approaches. *Journal of Cognitive Neuroscience*, **18**, 199-211.

[31] Schön, D., Magne, C., & Besson, M. (2004). The music of speech: music training facilitates pitch processing in both music and language. *Psychophysiology*, **41**, 341-349.

[32] Burnham, D., & Brooker, R. (2002) Absolute pitch and lexical tones: Tone perception by non-musician, musician, and absolute pitch non-tonal language speakers. In *7th International Conference on Spoken Language Processing*, Denver, USA, pp. 257-260.

[33] Schwanhäußer, B. & Burnham, D. (2005) Lexical tone and pitch perception in tone and non-tone language speakers. In *Proceedings of the 9th European Conference on Speech Communication and Technology*. Bonn, Germany, pp 1701-1704.

[34] Alexander, J. A., Wong, P. C. M. & Bradlow, A. R. (2005). Lexical tone perception in musicians and non-musicians. *Proceedings of Interspeech 2005*, Lisbon, Portugal.

[35] Bidelman, G.M., Gandour, J.T., & Krishnan, A. (2011). Cross-domain effects of music and language experience on the representation of pitch in the human auditory brainstem. *Journal of Cognitive Neuroscience*, **23**, 425-434.

[36] Delogu, F., Lampis, G. & Belardinelli, M. O. (2010). From melody to lexical tone: Musical ability enhances specific aspects of foreign language perception, *European Journal of Cognitive Psychology*, **22**, 46-61.

[37] Lee, C-Y. & Hung, T.-H. (2008). Identification of Mandarin tones by English-speaking musicians and nonmusicians. *Journal of the Acoustical Society of America*, **124**, 325-3248.

[38] Wong, P. C. M., Skoe, E., Russo, N. M., Dees, T., and Kraus, N. (2007). Musical experience shapes human brainstem encoding of linguistic pitch patterns. *Nat Neurosci,* **10**, 420-422.

[39] Moreno S., Marques C., Santos A., Santos M., Castro S., L. & Besson, M. (2009). Musical training influences linguistic abilities in 8-year-old children: More evidence for brain plasticity. *Cerebral Cortex*, **19**, 712-723.

[40] Marques C., Moreno S., Castro S. L., Besson, M. (2007). Musicians detect pitch violation in a foreign language better than nonmusicians: behavioral and electrophysiological evidence. *J Cogn Neurosci*., **19**, 1453-63.

[41] Stevens, C. J. (2012). Music perception and cognition: A review of recent cross-cultural research. *Topics in Cognitive Science*, **4**, 653-667

[42] Pfordresher, P. Q., and Brown, S. (2009). Enhanced production and perception of musical pitch in tone language speakers. *Attent Percep Psychophys*., **71**, 1385-1398.

[43] Hove, M. J., Sutherland, M. E. & Krumhansl, C. L. (2010). Ethnicity effects in relative pitch. *Psychonomic Bulletin & Review*, **17**, 310-316.

[44] Krishnan, A., Xu, Y., Gandour, J., & Cariani, P. (2005). Encoding of pitch in the human brainstem is sensitive to language experience. *Cognitive Brain Research*, **25**, 161-168.

[45] Stevens, C. J., Keller, P. E., & Tyler, M. D. (2013). Tonal language background and detecting pitch contour in spoken and musical items. *Psychology of Music*, **41**, 59-74.

[46] Bent, T., Bradlow, A. R. & Wright, B. A. (2006). The influence of linguistic experience on the cognitive processing of pitch in speech and nonspeech sounds. *Journal of Experimental Psychology: Human Perception and Performance*, **32**, 97-103.

[47] Deutsch, D., Dooley, K., Henthorn, T. and Head, B. (2009). Absolute pitch among students in an American music conservatory: Association with tone language fluency. *Journal of the Acoustical Society of America, April*, **125**, 2398-2403.

[48] Deutsch, D., Henthorn, T., Marvin, E., & Xu H-S. (2006). Absolute pitch among American and Chinese conservatory students: Prevalence differences, and evidence for a speech-related critical period. *Journal of the Acoustical Society of America*, **119**, 719-722.

[49] Hyde, K. L., and Peretz, I. (2004). Brains that are out of tune but in time. *Psychological Science* **15**, 356-360.

[50] Liu, F., Patel, A. D., Fourcin, A., et al (2010). Intonation processing in congenital amusia: discrimination, identification and imitation. *Brain*, **133**, 1682-93.

[51] Thompson, W. F., Marin, M. M., and Stewart, L. (2012). Reduced sensitivity to emotional prosody in congenital amusia rekindles the musical protolanguage hypothesis. *Proc Natl Acad Sci USA*, **109**, 19027-32.

[52] Nan, Y., Sun, Y., and Peretz, I. (2010). Congenital amusia in speakers of a tone language: association with lexical tone agnosia. *Brain*, **133**, 2635-42.

[53] Jiang, C, Hamm, J. P., Lim, V. K., et al (2010). Processing melodic contour and speech intonation in congenital amusics with Mandarin Chinese. Neuropsychologia, **48**, 2630-9.

[54] Jiang, C, Hamm, J. P., Lim, V. K., et al (2012). Impaired categorical perception of lexical tones in Mandarin-speaking congenital amusics. *Mem Cognit*, **40**, 1109-21.

[55] Nguyen, S., Tillmann, B., Gosselin, N., et al (2009). Tonal language processing in congenital amusia. *Ann N Y Acad Sci*, **1169**, 490-3.

[56] Tillmann, B, Burnham, D, Nguyen, S, et al (2011). Congenital Amusia (or tone-deafness) Interferes with pitch processing in tone languages. *Front Psychol*, **2**, 120.

[57] Liu, F, Jiang, C, Thompson, W. F., et al (2012). The mechanism of speech processing in congenital amusia: evidence from Mandarin speakers. *PLoS One*, **7**, e30374.

[58] Liu, F, Xu, Y, Patel, A. D., et al (2012). Differential recognition of pitch patterns in discrete and gliding stimuli in congenital amusia: evidence from Mandarin speakers. *Brain Cogn*, **79**, 209-15.

[59] Wong, P. C., Ciocca, V., Chan, A. H., et al (2012). Effects of culture on musical pitch perception. *PLoS One*, **7**, e33424.

[60] Jones, J. L., Lucker, J., Zalewski, C., et al (2009b). Phonological processing in adults with deficits in musical pitch recognition. *J Commun Disord.*, **42**, 226-34.

[61] Loui, P., Kroog, K., Zuk, J., et al (2011). Relating pitch awareness to phonemic awareness in children: implications for tone-deafness and dyslexia. *Front Psychol.*, **2**, 111.

[62] Wong, P. C. M., & Perrachione, T. K. (2007). Learning pitch patterns in lexical identification by native English-speaking adults. *Applied Psycholinguistics*, **28**, 565-585.

[63] Vroomen, J., and de Gelder, B. (2000). Sound enhances visual perception: cross-modal effects of auditory organization on vision. *J Exp Psychol Hum Percept Perform.*, **26**, 1583-1590.

[64] Caclin, A., Bouchet, P., Djoulah, F., Pirat, E., Pernier, J., and Giard, M. H. (2011). Auditory enhancement of visual perception at threshold depends on visual abilities. *Brain Res.*, **1396**, 35-44.

[65] Galvin, J. J., Fu, Q. J. & Nogaki, G. (2007). Melodic contour identification by cochlear implant listeners. *Ear & Hearing*, **28**, 301-319.

[66] Albouy, P., Leveque, Y., Hyde, K. L., Bouchet, P., Tillmann, B. & Caclin, A. (2014). Boosting pitch encoding with audiovisual interaction in congenital amusia. Manuscript submitted for publication.

[67] Thompson, W. F. & Balkwill, L-L. (2010). Cross-cultural similarities and differences. In Patrik Juslin and John Sloboda (Eds.), *Handbook of music and emotion: Theory, research, applications* (pp. 755-788). Oxford: Oxford University Press.