# Neural substrates of processing syntax and semantics in music

Stefan Koelsch

Junior Research Group Neurocognition of Music, Max Planck Institute for Human Cognitive and Brain Sciences, Leipzig, Germany.

Growing evidence indicates that syntax and semantics are the basic aspects of music. After the onset of a chord, initial music–syntactic processing can be observed at about 150-400 ms and processing of musical semantics at about 300-500 ms. Processing of musical syntax activates inferior frontolateral cortex, ventrolateral premotor cortex and presumably the anterior part of the superior temporal gyrus. These brain structures have been implicated in sequencing of complex auditory information, identification of structural relationships, and serial prediction. Processing of musical semantics appears to activate posterior temporal regions. The processes and brain structures involved in the perception of syntax and semantics in music have considerable overlap with those involved in language perception, underlining intimate links between music and language in the human brain.

### 9.1. Introduction

Music is one of the oldest, and most basic, socio-cognitive domains of the human species. Primate vocalizations are mainly determined by music-like features (such as pitch, amplitude and frequency modulations, timbre and rhythm), and it is assumed that human musical abilities played a key phylogenetical part in the evolution of language (Zatorre and Peretz 2003). Likewise, it is assumed that, ontogenetically, infants' first steps into language are based on prosodic information, and that musical communication in early childhood (such as maternal music) has a major role for emotional, cognitive and social development of children (Trehub 2003). The music faculty is in some respects unique to the human species; only humans compose music, learn to play musical instruments and play instruments cooperatively together in groups. Playing a musical instrument in a group is a tremendously demanding task for the human brain that potentially engages all cognitive processes that we are aware of. It involves perception, action, learning, memory, and emotion, making music an ideal tool to investigate human cognition and the underlying brain mechanisms. The relatively young discipline of 'neurocognition of music' includes a wide field of biopsychological research, beginning with the investigation of psychoacoustics and the neural coding of sounds, and ending with brain functions underlying cognition and emotion during the perception and production of highly complex musical information (Zatorre and Peretz 2003, Avanzini et al. 2003). This review focuses on two basic dimensions of music perception.

# 9.2. The processing of syntax and of semantics in music

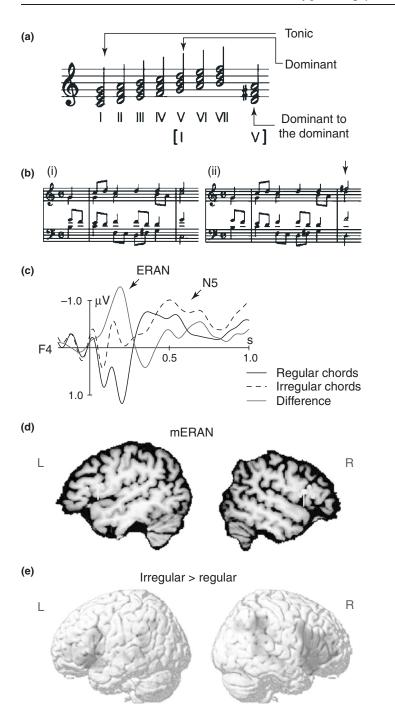
All types of music are guided by certain regularities. These regularities constrain, for example, how individual tones, simultaneous tones (i.e. intervals and chords) and durations of tones are arranged to form meaningful musical phrases. Obviously, many regularities are culturespecific and differ between musical styles. To date, the processing of regularities has mainly been investigated with respect to major–minor tonal music; this music is formed on the basis of the major–minor tonal system that underlies the majority of Western music. Basic principles and regularities of this tonal system have been described in music theory. One aspect of these regularities build melodic, rhythmic and metric structure). The regularity-based arrangement of chord functions builds a harmonic structure, and might be regarded as part of a major–minor tonal syntax (Fig. 1a, b). Listeners who grew up in a Western culture are usually quite familiar with these regularities (even if they have not received formal musical training), presumably because of their listening experiences in everyday life. It is unknown if listeners who are completely unfamiliar with Western tonal music can also recognize basic syntactic irregularities of major–minor tonal music.

### 9.2.1 Brain indices of processing harmonic structure

Processing of chord functions has been investigated behaviorally (Tillmann et al. 2000), and with neurophysiological measures such as electroencephalography (EEG) (Janata 1995, Patel

#### Fig. 1:

Neural correlates of music-syntactic processing. (a) In major-minor tonal music, chord functions are arranged within harmonic sequences according to certain regularities. Chord functions are the chords built on the tones of a scale. The chord on the first scale tone, for example, is denoted as the tonic and the chord on the fifth scale tone as the dominant. The major chord on the second tone of a major scale can be interpreted as the dominant to the dominant (square brackets). (b) One example for a regularity-based arrangement of chord functions is that the dominant-tonic progression is a prominent marker for the end of a harmonic sequence, whereas a tonic-dominant progression is unacceptable as a marker of the end of a harmonic sequence. (i) The sequence shown ends on a regular dominant-tonic progression, (ii) the final chord of this sequence is a dominant to the dominant (arrow). This chord function is irregular, especially at the end of a harmonic progression (sound examples are available at www.stefan-koelsch.de/TC\_DD). (c) Electric brain potentials (in mV) elicited by the final chords of the two sequence types presented in b (recorded from a right-frontal electrode site [F4] from twelve subjects). Both sequence types were presented in pseudorandom order with equal probability in all twelve major keys. Brain responses to irregular chords clearly differ from those to regular chords. The first difference between the two black waveforms is maximal at about 0.2 s after the onset of the chord (this is best seen in the red difference wave, which represents regular subtracted from irregular chords) and has a right-frontal preponderance. This early right anterior negativity (ERAN) is usually followed by a later negativity, the N5 (short arrow). (d) With MEG, the magnetic equivalent of the ERAN was localized to the inferior frontolateral cortex (adapted from Maess et al. with permission of Nature Publishing Group [http://www.nature.com/] (Maess et al. 2001); single-subject dipole solutions are indicated by blue disks, yellow dipoles indicate the grand-average of these source reconstructions). (e) fMRI data obtained from 20 subjects using a similar chord-sequence paradigm (the statistical parametric maps show areas that are more strongly activated during the processing of irregular than during the processing of regular chords). Corroborating the MEG data, the fMRI data indicate activations of IFLC. Additionally, the fMRI data indicate the activations of the ventrolateral premotor cortex, the anterior portion of the STG, and the posterior temporal lobe structures. Current Opinion in Neurobiology 2005, 15, 207-212.



et al. 1998, Regnault et al. 2001), magnetoencephalography (MEG) (Maess et al. 2001) and functional magnetic resonance imaging (fMRI) (Tillmann et al. 2003, Koelsch et al. 2005). These studies used chord sequence paradigms in which chords presented at particular positions within harmonic sequences are structurally more or less (ir)regular. Figure 1b shows musical sequences ending on music-syntactically regular and irregular chord functions (see Fig. 1 legend for details). Note that the final chord of the irregularGo! sequence (right panel of Fig. 1b) does not represent a physical irregularity. It is, thus, not possible to detect the irregular chords on the basis of the operations of cognitive modules that detect physical irregularities (such as the auditory sensory memory). It is only possible to detect the irregular chords on the basis of the operations of a cognitive module that serves the processing of musical structure.

The event-related brain potential (ERP) data of Fig. 1c illustrate the time course of activity of this module. Music-syntactically irregular chords elicit an early right anterior negativity (ERAN): this ERP effect is often maximal at about 200 ms after the onset of the chord, and is strongest over right-frontal electrode leads (although the ERAN is also clearly present over the left hemisphere, and is sometimes bilateral). Interestingly, the ERAN can be elicited in participants without formal musical training. That is, even if participants are not familiar with concepts such as 'tonic', or 'dominant', their brains have a sophisticated (implicit) knowledge about major-minor tonal syntax, and process this musical information surprisingly rapidly and accurately according to this knowledge. These findings are in line with several studies indicating that the ability to acquire knowledge about musical regularities effortlessly, and the ability to process musical information skillfully according to this knowledge, is a general ability of the human brain (details have been reviewed elsewhere (Avanzini et al. 2003, Tillmann et al. 2000)). This general ability underscores the biological relevance of music. Musical abilities are important, for example, with regard to language perception: in tonal languages, changes in pitch lead to changes in word meaning, and in both tonal and non-tonal languages prosody (i.e. the musical features of language such as melody, timbre, rhythm and metre) is of crucial importance for the coding of both the structure and the meaning of speech. Corroboratively, recent EEG studies revealed similarities for the processing of intonational phase boundaries in language and music (Knoesche et al. 2005, Steinhauer et al. 1999) and showed that musical training can facilitate the processing of pitch contour in spoken (nontonal) language (Schon et al. 2004).

The neural mechanisms underlying the generation of the ERAN can operate independently of attention (although the amplitude of the ERAN is influenced by attentional demands): the ERAN can be elicited when subjects read a book, play a video game or are lightly sedated with propofol (Heinke et al. 2004). The ERAN is sensitive to musical training (Koelsch et al. 2002b), and can be elicited in children aged 5 years and perhaps even earlier (Koelsch et al. 2003).

### 9.2.2 Spatial aspects of music-syntactic processing

Using MEG, it was found that the processing of music syntactically irregular chords activates the inferior part of Brodmann's area (BA) 44, that is, the inferior frontolateral cortex

(IFLC; (Maess et al. 2001), see Fig. 1d). This area in the left hemisphere is often denoted as 'Broca's area', an area that has also been implicated in the processing of linguistic syntax. With fMRI, it has been demonstrated that the processing of unexpected chords does not only activate Broca's area (and the homotope area in the right hemisphere) (Tillmann et al. 2003, Koelsch et al. 2002a, 2005) but also activate posterior temporal regions (Koelsch et al. 2002a, 2005) (see also Fig. 1e). Both Broca's area and posterior temporal regions are crucially involved in the processing of language (Friederici 2002); the interplay between these structures has for a long time been thought to be language-specific. The data presented in Fig. 1e demonstrate that the cortical 'language-network' is also involved in the processing of music. This network often shows a right-hemispheric weighting in the musical domain, and a left-hemispheric weighting in the language domain (specializations of the two hemispheres for different features of auditory information have been discussed elsewhere (Zatorre et al. 2002)).

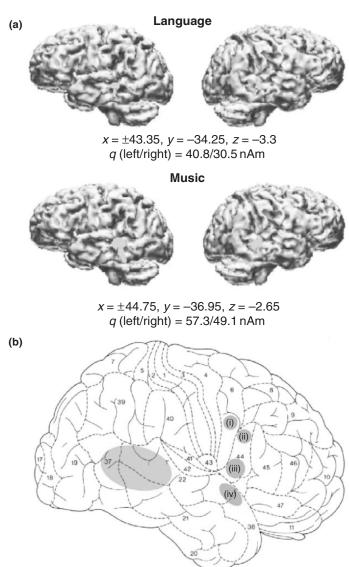
The ERAN is reminiscent of early anterior negativities that correlate with the early detection of an error in the syntactic structure of a sentence (usually observed with a maximum over the left hemisphere). The early left anterior negativity (ELAN), for example, has been observed in response to words with unexpected syntactic properties in sentences (phrase structure violations) (Friederici 2002). That is, both ERAN and ELAN are sensitive to violations of an expected structure. Moreover, the generation of the ELAN appears to rely on neuronal generators that overlap with those of the ERAN, in that both components receive contributions from the same brain region in the inferior frontolateral cortex (lower part of BA 44) (Maess et al. 2001), and possibly from the anterior superior temporal gyrus (aSTG; see also below). Taken together, these findings indicate a noticeable overlap of neural resources that are engaged for the (early) processing of syntax in music, and syntax in language.

Besides IFLC, two additional structures have been observed in relation to music-syntactic processing: the ventrolateral premotor cortex (vlPMC) and the aSTG (Figs. 1, 1e, 2, 2b). Activations of IFLC along with the aSTG have been reported in previous functional imaging studies on syntactic processing using musical (Koelsch et al. 2002a, 2005) and linguistic stimuli (Friederici 2002, Friederici et al. 2000, 2003).

Activations of the IFLC (BA44), often along with the vlPMC, have been reported by a number of functional imaging studies using musical stimuli, linguistic stimuli, auditory oddball paradigms, pitch discrimination tasks, and serial prediction tasks (Tillmann et al. 2003, Koelsch et al. 2002a, 2005, Friederici 2002, Janata et al. 2002, Doeller et al. 2003, Gaab et al. 2003, Schubotz and von Cramon 2002). On a more abstract level, the IFLC (BA44) and the vlPMC have been implicated in the analysis, recognition and prediction of sequential auditory information (Schubotz and von Cramon 2002). Conway and Christiansen 2001, Huettel et al. 2002, Janata and Grafton 2003). Frontoopercular cortex (along with vlPMC) identifies structural relationships (rather than simple acoustic properties) among events occurring within auditory sequences, and these areas are involved in a fast short-term prediction of upcoming events; violations of predictions activate these areas (Schubotz and von Cramon 2002). The presentation of an irregular chord function violates the expectancies

#### Fig. 2:

Spatial aspects of processing syntax and semantics in music. (a) Neural generators of the N400 effect elicited by target words that were semantically (un)related to preceding sentences (top, blue dipoles), or musical excerpts (bottom, brown dipoles). The topology of the neural sources of the N400 did not differ between the language and the music condition (adapted from (Koelsch et al. 2004); x-, y-, and z-coordinates refer to standard stereotaxic space, dipole moments [q] are given in nanoamperemeters). (b) Spatial aspects of processing syntax and semantics in music. Orange areas represent activation foci described in previous imaging studies on musicsyntactic processing: (i) vIPMC, IFLC ([ii] superior and [iii] inferior pars opercularis) and (iv) anterior STG. The blue oval represents an area that is assumed to be involved in the processing of musical semantics, and in the integration of semantic and syntactic information.



of listeners familiar with the regularities of tonal music. Unusual calculations of the relationship between the irregular chord function and the preceding harmonic context presumably activate a network comprising the pars opercularis in the IFLC, the vlPMC and presumably the anterior STG. These calculations are related to the sequencing of the chords, and the detection of a violation of a serial prediction. Whether neural substrates of these processes can functionally and anatomically be distinguished from each other remains to be specified. Likewise, it is not known if the representation of musical syntax is located in the same areas that are involved in the processing of musical syntax (Patel 2003).

It has been suggested that there might be an immediate link between the prediction of upcoming events and the representation of corresponding motor schemas in the lateral premotor cortex (PMC) that enables an immediate mapping of perception onto action, that is, premotor programs for articulation, or 'vocal plans' (Schubotz and von Cramon 2002). Such an aping is needed, for example, when singing along in a group, and is presumably also involved in the learning and understanding of musical syntax.

The ERAN is not the only electrophysiological index of music-syntactic processing, ERP studies investigating the processing of musical structure report a variety of ERP components such as P300 (Janata 1995), RATN (right anterior temporal negativity) (Patel et al. 1998), and P600 (Patel et al. 1998, Regnault et al. 2001, Besson and Faita 1995). The P600 (a positivity maximal at about 600 ms) appears to be related to processes of structural integration during the perception of music or language. Because the P600 can be observed in response to structural incongruities of both music and language (Patel et al. 1998), it has been suggested that resources for syntactic integration are shared between music and language (Patel 2003). The similarities between ERAN and ELAN suggest that not only neural resources for late but also for earlier syntactic processes are shared between music and language.

#### 9.3. Processing meaning in music

Music can transfer meaningful information, and is an important means of communication. Most theorists distinguish between different aspects of musical meaning: first, meaning that emerges from common patterns or forms (e.g. musical sound patterns that resemble sounds of objects, or qualities of objects), second, meaning that arises from the suggestion of a particular mood (e.g. happy), third, meaning inferred by extramusical associations (e.g. any national anthem), and fourth, meaning that emerges from combinations of formal structures that create tension (e.g. when perceiving an unexpected chord) and resolution (Meyer 1956). The emergence of meaning based on the processing of musical structure requires the integration of both expected and unexpected events into a larger, meaningful musical context.

Such processes of musical integration appear to be reflected in a later negativity evoked by unexpected (irregular) chord functions (Fig. 1b, c). This negativity usually reaches its maximal amplitude at about 500 ms after the onset of a chord and has been denoted as N5 (Koelsch et al. 2000; Fig. 1c). Note that the processes of semantic integration during the perception of language are reflected in the N400 (Kutas and Federmeier 2000), a negativity peaking at about 400 ms after the onset of a word. Similarly to the N400 amplitude, which correlates with the amount of semantic integration required by a word, the N5 amplitude is related to the amount of harmonic integration required by a musical event (Koelsch et al. 2000).

Differences in scalp topography between N400 and N5 indicate that these two ERP components do not reflect identical cortical processes. However, because the N5 roughly resembles the N400, and because the cognitive processes following musical expectancy violations have theoretically been related to the processing of meaningful information, it

appears likely that the N5 reflects neural operations that are at least partly related to the processing of musical meaning, and that the N5 entails processes that might also contribute to the generation of the N400 (note that irregular chord functions, and deceptive cadences, are prominent elements of major-minor tonal music that are used by composers as a means of expression). The N5 has not been localized to date, but it is possible that the N5 receives contributions from those posterior temporal lobe structures that have been shown with fMRI to be activated during the processing of unexpected chords. These structures are also known to be involved in the processing of lexical-semantic aspects, that is, meaning of language (Friederici 2002).

The N400 has recently been used to investigate the processing of musical semantics in a semantic priming paradigm (Koelsch et al. 2004). In this study, sentences and musical excerpts were presented as prime stimuli. The prime stimuli were semantically either related or unrelated to a target word that followed the prime stimulus. For example, the sentence 'The gaze wandered into the distance' primes the word 'wideness' (semantically related), rather than the word 'narrowness' (semantically unrelated). Analogously, certain musical passages, for example, those from Mozart's symphonies, prime the word 'angel', rather than the word 'scallywag'.

In the language condition (i.e. when target words followed the presentation of sentences), unrelated words elicited a clear N400 effect (this is a classical semantic priming effect). This semantic priming effect was also observed when target words followed musical excerpts. That is, target words that were semantically unrelated to a preceding musical excerpt also elicited a clear N400. The N400 effects did not differ between the language condition (in which the target words followed sentences) and the music condition (in which the target words followed sentences) and the music condition (in which the target words followed musical excerpts), neither with respect to amplitude nor with respect to latency or scalp distribution. Figure 2a shows the results of a source analysis of the N400 effects. In both conditions, the main sources of these effects were localized bilaterally in the posterior part of the medial temporal gyrus (BA 21/37), in proximity to the superior temporal sulcus. As mentioned above, these regions have been implicated in the processing of semantic information during language processing (Friederici 2002, Baumgaertner et al. 2002).

The N400 effect in the music condition demonstrates that musical information can have a systematic influence on the semantic processing of words. The N400 effects did not differ between the music and the language condition, indicating that musical and linguistic priming can have the same effects on the semantic processing of words. That is, the data demonstrate that music can activate representations of meaningful concepts (and that, thus, music is capable of transferring considerably more meaningful information than previously believed), and that the cognitive operations that decode meaningful information while listening to music can be identical to those that serve semantic processing during language perception.

The N400 effect was observed for both abstract and concrete words, showing that music can convey both abstract and concrete semantic information. Moreover, effects were also observed when emotional relationships between prime and target words were balanced, indicating that music does not only transfer emotional information.

# 9.4. Conclusions

The present findings provide information about the processing of musical syntax and musical semantics. Results indicate that the human brain processes music and language with overlapping cognitive mechanisms, in overlapping cerebral structures. This view corresponds with the assumption that music and speech are intimately connected in early life, that musical elements pave the way to linguistic capacities earlier than phonetic elements, and that melodic aspects of adult speech to infants represent the infants' earliest associations between sound patterns and meaning (Fernald 1989), and between sound patterns and syntactic structure (Jusczyk and Krumhansl 1993). Thus, despite the view of some linguists that music and language are strictly separate domains (Pinker 1997), the combined findings indicate that the human brain engages a variety of neural mechanisms for the processing of both music and language, underscoring the intimate relationship between music and language in the human brain.

### References

Avanzini G, Faienza C, Miniacchi D, Lopez L, Majno M (2003) The Neurosciences and Music. Annals of the New York Academy of Sciences 999. This book provides another extensive overview of the different fields investigated within neurocognition of music

Baumgaertner A, Weiller C, Büchel C (2002) Event-related fMRI reveals cortical sites involved in contextual sentence integration. Neuroimage 16, 736-745

Besson M, Faita F (1995) An event-related potential study of musical expectancy: Comparison of musicians with nonmusicians. J Exp Psych: Hum Perc Perf 21, 1278-1296

Conway CM, Christiansen MH (2001) Sequential learning in nonhuman primates. Trends Cogn Sci 5, 539-546

Doeller CF, Opitz B, Mecklinger A, Krick C, Reith W, Schroger E (2003) Prefrontal cortex involvement in preattentive auditory deviance detection: neuroimaging and electrophysiological evidence. NeuroImage 20, 1270-1282

Fernald A (1989) Intonation and communicative intent in mothers' speech to infants: is the melody the message? Child Dev 60, 1497-1510

Friederici AD (2002) Towards a neural basis of auditory sentence processing. Trends Cogn Sci 6, 78-84

Friederici AD, Wang Y, Herrmann CS, Maess B, Oertel U (2000) Localisation of early syntactic processes in frontal and temporal cortical areas: An MEG study. Hum Brain Mapp 11, 1-11

Friederici AD, Rüschemeyer S, Hahne A, Fiebach CJ (2003) The role of left inferior frontal and superior temporal cortex in sentence comprehension: localizing syntactic and semantic processes. Cereb Cortex 13, 170-177

Gaab N, Gaser C, Zaehle T, Jancke L, Schlaug G (2003) Functional anatomy of pitch memory – an fMRI study with sparse temporal sampling. Neuroimage 19, 1417-1426

Heinke W, Kenntner R, Gunter TC, Sammler D, Olthoff D, Koelsch S (2004) Differential effects of increasing propofol sedation on frontal and temporal cortices: an ERP study. Anesthesiology 100, 617-625. Using ERPs and music, the authors investigated the effects of sedative drugs (propofol) on auditory processing. It was found that propofol affects at lower doses processing of 'higher' cognitive functions located in multimodal cortex, whereas functions located in primary auditory cortical areas remain unaffected

Huettel SA, Mack PB, McCarthy G (2002) Perceiving patterns in random series: dynamic processing of sequence in prefrontal cortex. Nat Neurosci 5, 485-490

Janata P (1995) ERP measures assay the degree of expectancy violation of harmonic contexts in music. J Cogn Neurosci 7, 153-164

Janata P, Birk JL, Van Horn JD, Leman M, Tillmann B, Bharucha JJ (2002) The cortical topography of tonal structures underlying Western music. Science 298, 2167-2170

Janata P, Grafton ST (2003) Swinging in the brain: shared neural substrates for behaviors related to sequencing and music. Nat Neurosci 6, 682-687

Jusczyk PW, Krumhansl CL (1993) Pitch and rhythmic patterns affecting infants' sensitivity to musical phrase structure. J Exp Psychol Hum Percept Perform 19, 627-640

Knoesche TR, Neuhaus C, Haueisen J, Alter K, Maess B, Witte OW, Friederici AD (2005) The Perception of Phrase Structure in Music. Human Brain Mapp (in press)

Koelsch S, Gunter TC, Friederici AD, Schroger E (2000) Brain indices of music processing: 'non-musicians' are musical. J Cog Neurosci 12, 520-541

Koelsch S, Gunter TC, v Cramon DY, Zysset S, Lohmann G, Friederici AD (2002a) Bach speaks: a cortical "language-network" serves the processing of music. Neuroimage 17, 956-966

Koelsch S, Schmidt BH, Kansok J (2002b) Influences of musical expertise on the ERAN: An ERP-study. Psychophysiology 39, 657-663

Koelsch S, Grossmann T, Gunter TC, Hahne A, Friederici AD (2003) Children processing music: electric brain responses reveal musical competence and gender differences. J Cogn Neurosci 15, 683-693

Koelsch S, Kasper E, Sammler D, Schulze K, Gunter TC, Friederici AD (2004) Music, language, and meaning: brain signatures of semantic processing. Nat Neurosci 7, 302-307. Using ERPs and a priming paradigm the authors show that music can activate representations of meaningful concepts. This is the first study to show that music can transfer meaningful information, and that this transfer relies on those neurophysiological processes engaged for the processing of semantics in language

Koelsch S, Fritz T, Schulze K, Alsop D, Schlaug G (2005) Adults and Children Processing Music: An fMRI Study. NeuroImage (in press)

Kutas M, Federmeier KD (2000) Electrophysiology reveals semantic memory use in language comprehension. Trends Cogn Sci 4, 463-470

Maess B, Koelsch S, Gunter TC, Friederici AD (2001) 'Musical syntax' is processed in Broca's area: an MEGstudy. Nat Neurosci 4, 540-545

Meyer LB (1956) Emotion and Meaning in Music. University of Chicago Press, Chicago

Patel A (2003) Language, music, syntax and the brain. Nat Neurosci 6, 674-681

Patel AD, Gibson E, Ratner J, Besson M, Holcomb PJ (1998) Processing syntactic relations in language and music: an event-related potential study. J Cogn Neurosci 10, 717-733

Pinker S (1997) How the mind works. Norton

**Regnault P, Bigand E, Besson M (2001)** Different brain mechanisms mediate sensitivity to sensory consonance and harmonic context: evidence from auditory event-related brain potentials. J Cogn Neurosci 13, 241-255

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

Schubotz Rl, von Cramon DY (2002) Predicting perceptual events activates corresponding motor schemes in lateral premotor cortex: an fMRI study. Neuroimage 15(4), 787-796

Steinhauer K, Alter K, Friederici AD (1999) Brain potentials indicate immediate use of prosodic cues in natural speech processing. Nat Neurosci 2, 191-196

Tillmann B, Bharucha J, Bigand E (2000) Implicit learning of tonality: a self-organized approach. Psychol Rev 107, 885-913

Tillmann B, Janata P, Bharucha JJ (2003) Activation of the inferior frontal cortex in musical priming. Brain Res Cogn Brain Res 16, 145-161

Trehub S (2003) The developmental origins of musicality. Nat Neurosci 6, 669-673

Zatorre RJ, Belin P, Penhune VB (2002) Structure and function of auditory cortex: music and speech. Trends Cogn Sci 6, 37-46

Zatorre RJ, Peretz I (2003) The Cognitive Neuroscience of Music. Oxford University Press. This book provides an extensive overview of the different fields investigated within neurocognition of music