Musica, Maestro!

Daniele Schön and Mireille Besson

Center for Research in Cognitive Neuroscience

CNRS-Marseille, France

Address Correspondence to : Mireille Besson
Centre de Recherche en Neurosciences Cognitives
CNRS-CRNC
31- Chemin Joseph Aiguier
13402 – Marseille Cedex 20
France
mbesson@lnf.cnrs-mrs.fr
(33) 491 164 305 (tel)
(33) 491 774 969 (fax)

> If you can walk you can dance. If you can talk you can sing.
>
> Zimbabwe proverb

1. Introduction

The series of experiments that we describe in this chapter are centered on the question of the specificity of the computations involved in language processing. In order to address this question, language is compared with another well-organized cognitive function, that, although very different in many respects, nevertheless presents interesting similarities with language : music. Language and music differ by the nature of their respective constitutive elements, phonemes, morphemes, words in spoken language and notes and chords in music, and their structural organization. While language is organized horizontally, a succession of sounds in time, music offers both a horizontal organization, the melody formed by the pitch relationship between successive sounds, and a vertical organization, the simultaneous production of two or more sounds such as in chords. Furthermore, language and music differ by their social function. While most authors will agree that the basic function of language is to express thoughts in order to communicate with other individuals, the function of music is a subject of controversies, may be because music is more culturally bound than language. In Western cultures, music has evolved in such a way as to become more and more isolated from other expressive forms. By contrast, in other cultures, in which the magical thought is still alive, the bound between music, song, dance, poetry and rite has not been lost (Schön and Schön, 1999; Sloboda, 2000; Blacking 1973). Ethnomusicological studies have often emphasized that music translates emotional experiences into artistic forms, and allows to communicate with the unknown, as can be seen from the religious rituals in primal tribes (e.g., Nadel, 1930; Von Hornbostel, 1928). Moreover, Kubik (1969) has pointed out that, in African cultures, music is the acoustic result of an action and action is an intrinsic part of musical performance. Motor patterns are themselves sources of aesthetic pleasure, independently from the sound they are associated with. This strong intertwining between music and action is even reflected in language, the same word being used in several African languages to refer to music and dance. This conception stands in marked contrast with the passive perceptive status of music within the Western tradition. Thus, for instance, most of the studies in the psychology of music are centered on music perception (and our studies are, unfortunately, of no exception), and music performance has received much less attention (but see Sloboda, 2000).

Aside from these differences, language and music have both been developed by all human cultures and seem specific to humans; they both ensure the cohesion of the social group and, as such, play a powerful social function in all human societies (**Arom & Khalfa**, 1999; Boucourechliev, 1993; Levman, 1992; Nadel, 1930). Blacking (1973) for instance, defines music as ‘sound that is organized into socially accepted patterns’. Moreover, he argues that every piece of music has its own inherent logic, as the creation of an individual reared in a particular cultural background. However, his claim that patterns of sounds reflect patterns of social organization seems somewhat coarse, and in need of further elaboration. Still, in much the same way that a context-sensitive grammar is a more powerful analytical tool than a context-free grammar, the cognitive systems underlying different styles of music shall be better understood if music is considered in context. Different musical styles should thus not be considered as ‘sonic objects’ but as humanly organized sound, whose patterns are related to the social and cognitive processes of a particular society and culture. Finally, both language and music rely on a sequential organization of sounds that unfold in time. These sound are in both cases characterized by their pitch, duration, intensity and timber. They are structured into separate units by variations in voice intonation, prosody, that tend to go down at the end of sentences, and by the cadence at the end of musical phrases (i.e., the classic

succession of tonic, dominant, sub-dominant and tonic chords at the end of musical phrases). Most importantly, both language and music rely on different levels of processing. In the first part of this chapter, we will lay the cognitive basis to compare language and music by describing the different levels of processing at play within the musical and linguistic systems. We will then quickly review some of the theoretical and experimental arguments in favor or against the question of the specificity of the computations involved in language processing. We will finally report some experiments conducted in the fields of neuropsychology and cognitive neurosciences, using brain imaging methods, and aimed at specifying the brain mechanisms and the cerebral structures involved in some aspects of language and music processing, before summarizing the results of our own research.

2. Music and language structure

From a cognitive perspective, music and language cannot be considered as single entities. In order to be analyzed and compared they need to be reduced to their constitutive elements. Within music, one classically differentiate the temporal (meter and rhythm), melodic (contour, pitch and interval) and harmonic (chords, voices) aspects. Each aspect most likely involves different types of processing, so that the processes called into play to process rhythm may differ from those involved in the processing of pitch and melodic intervals. Similarly, within language, at least four different levels of processing have been taken into consideration. The phonetic-phonological level, that comprises both segmental (phonemes) and suprasegmental (prosody); the morpho-syntactic level, that encompasses the combination of phonemes into morphemes and of morphemes into words; the syntactic level, that governs the relations between words, and the lexico-semantic level, with access to the meaning of words and sentences. Finally, while often ignored in psycholinguistic and neurolinguistic experiments, the pragmatic level that comprises discourse organization and contextual influences, represents an essential aspect of language organization. While it is of great interest to compare these different levels of processing in language and music, one should keep in mind that this comparison might highlight either their similarities or their differences depending upon the grain chosen for the analysis. Thus, similarities at one level of processing may be interpreted as differences at another level.

2.1. Phonetic/phonological level

At a phonetic level two interesting phenomena have been demonstrated in music as well as in language: categorical perception and phonemic restoration (Aiello, 1994). The categorical perception phenomenon occurs when discrimination within the same category is poor, but discrimination between different categories is good. The segmentation of the sound continuum into discrete units (pitches or phonemes) is found in all music and languages. In Western music, sounds are categorized according to tonal parameters, thus into the intervals of a scale. Luckily for the violin player, while small errors are often detected by musicians, they are not heard by naïve, non-musician listeners who show a large range of acceptability (Locke and Kellar, 1973). This is similar in a way to an English speaker going to Nepal, and trying to detect the differences between some phonemes of the Nepali alphabet like /ka/, /kha/ and /khha/. Moreover, as speakers of different languages have a categorical perception of the phonemes of their own language, musicians of different musical cultures are sensitive to different pitch changes, depending upon the musical system they are familiar with. Finally, in playing music as well as in speaking language, categorical perception seems to favor both a better and easier recognition and comprehension, regardless of individual differences.

However, when comparing phonemes and intervals of a musical scale, several differences must also be taken into consideration. While the variability of the number of pitches by octave across musical cultures is relatively small, the number of phonemes largely differ between languages (from 11, in Polynesian, to 141 in the language of the Bushmen; Pinker, 1994), with 44 phonemes in English and 36 in French). Moreover, some of the perceptual properties of the basic elements in music have no equivalent in language, as for instance the fact that octaves are perceived as equivalent in almost all cultures. This effect is linked with the fact that two notes separated by an octave are related by a simple frequency ratios of 2/1. Generally speaking, the relationships between different pitches in a musical piece are much simpler than the relationships between different phonemes in a linguistic sentence.

The second phenomenon is that of phonemic restoration, occurring in both music and language. It occurs when a noise, replacing a pitch in music or a phoneme in language, is not perceived by the listener, who believe they heard an intact auditory signal. In this case, lexical knowledge and lexico-semantic expectations in one case, and musical knowledge and expectancies in the other case, take over low level processing analysis, filling in the missing information.

2.2. Syntactic level

All languages are organized according to a syntactic structure that some authors consider to be universal (Chomsky 1988, 1991). Thus, verbs and nouns are always present. However, the order in which those elements are presented vary between languages (Subject – Verb – Object (French, English, …); Subject – Object – Verb (German; Japanese, ….); Verb – Object – Subject (Malgache, …) (François, 1988). Even if it is common to refer to a musical syntax or a musical grammar, the extent to which this analogy goes beyond a simple metaphor remains to be determined. Music perception shares universal laws of auditory perception that have a strong influence on music syntactic structure. For instance, the perception of a musical phrase is automatically influenced by factors such as the grouping of discrete notes into meaningful sequences. Halpern and Bower (1982) showed that musical patterns are memorized by musicians as single chunks, while the recall of random or less frequently used patterns is more likely to rely on unspecific strategies, similar to those used by non-musicians. Different groups or chunk of notes constitute a musical phrase, with precise boundaries, just like in language. Several studies, conducted on music reading, have demonstrated the extent to which musicians rely on significant structures, ignore superfluous material, and fail to notice printing errors (Sloboda, 1985). Moreover, the feeling of closure elicited by a cadence at the end of a musical phrase is similar to the Gestalt principle of closure in visual perception. Some musical phenomena such as grouping are universally shared, and others, just as verbal language, are culturally shared. However, even if there is such a thing as a musical grammar, the rules seem more flexible and ambiguous than the syntactic rules used in language. If the communicative function of language favors syntactical stability, ambiguity is a key element of the grammar and aesthetics of music (Aiello, 1994). There are always several ways to perceive and enjoy a musical piece. Finally, musical elements are most often played simultaneously and each element may have its own “syntax”. As noted previously, this vertical dimension of musical structure, commonly referred to as harmony, is not present in language. While different words sung at the same time may melt in a sublime combination of rhythm, melody and harmony (as in the polyphonic compositions of Gesualdo da Venosa, 1560-1613), different words produced at the same time by different speakers will only create an unpleasant cacophony, like in a political debate!

2.3. Meaning, expectancy and affect

Even if the similarities and differences between music and language depend upon the level of details considered for the analysis, one fundamental difference nevertheless remains. While the meaning of words is understood in relation to an extra-linguistic designated space, music is considered mostly self-referential (Jakobson, 1973; Meyer, 1956; Kivy, 1991; Boucourechliev, 1993). This does not mean that music is asymbolic. However, while the meaning of words is defined by an arbitrary convention relating sounds to meaning, notes or

chords have no extra-musical space in which they would acquire meaning. The internal sense of music may be conceived as something that goes beyond any objective reference structure and the possibilities of verbal language (Piana, 1991). Much as Wittgenstein (1953) who asked: “Describe the coffee aroma!”, music is the kingdom of the ineffable. As stated by Leonard Meyer in his wonderful book "*Emotion and meaning in music*" (1956): “Music means itself. That is, one musical event (be it a tone, a phrase or a whole section) has meaning because it points to and makes us expect another musical event” (p.35). Interestingly, this statement not only highlights one of the most important differences between language and music, that is the unsolved question of musical (a)semantics, but also emphasizes one of their strongest similarity: in their own way, both systems generate strong expectancies. Just as a specific word is expected within a specific linguistic context, specific notes or chords are expected at a given moment within a musical phrase. Either these expectations are fulfilled giving rise to resolution or satisfaction, or they are not fulfilled, giving rise to tension or surprise.

An important consequence of the “expectation-realization theory” proposed by Meyer (ib.) is its relation to musical affect. However, one problem with a theory of musical affect based upon the intuitive notion of expectation is that when listening to a well-known piece, one can anticipate how the music will continue. As a consequence, the affect evoked by the piece should diminish proportionally to the degree of familiarity. This is not the case and Jackendoff (1991) suggests a potential explanation. First, he argues that we should not believe that expectation may ‘bear the entire burden of deriving affect’ (Jackendoff, 1991). Others factors such as tempo, intensity, non-mechanical interpretation of music also influence musical emotions. Still, the structure of music has intrinsic points of instability that tend to resolve, and the tension / resolution phenomenon induces affects. Moreover, tensions are perceived at different levels depending upon the analysis performed. In analogy to language, the author (Jackendoff, 1991) points out that a modular and informationally encapsulated parser might be at work, independently from conscious memory. This independence from memory may explain why we keep enjoying a piece on repeated hearings: an autonomous parser will keep analyzing and recreating whatever structure is retrieved from memory. Then 'surprise will still occur within the parser' (Jackendoff, 1991).

Do these considerations really lead to a better insight on the frequently asked question 'what is the meaning of music?'. We may suggest that, rather than asking this question, a more fruitful approach is to reflect on 'what can I do with sounds?'. We may then discover that music is first of all a set of choices and the flow of these choices might possibly become visible as a musical thought. Behind all this, the image of children playing appears. When the child plays with small wood blocks we could say that the game is a way of answering the question: "what can I do with my small wood blocks?". From the pleasure of playing we get directly into the aesthetical pleasure. Then we could possibly describe music as a game: the musician plays an instrument, the composer plays with sounds and the listener plays with his perception and emotions (Leipp, 1977). As every game, music obeys a set of rules, even if they may vary depending upon the period, the history, the culture.

Concluding this quick and necessarily incomplete excursus, and before turning to language specificity, it is important to keep in mind that musical meaning is the sum of analytic approaches (musical parser), individual and/or cultural associations to the external/internal world (during some periods in the last centuries "music was conceived as conveying precise emotional and conceptual meanings, established by codes, or at least, *repertoires*"; Eco, 1979, pp. 11) and aesthetic reaction. The importance of the aesthetic component of music becomes evident in considering that "the form of a work of art gains its aesthetics validity precisely in proportion to the number of different perspectives from which it can be viewed and understood" (Eco, 1989, pp, 3)

3. Specificity of the computations involved in language processing?

3.1. The Generative Grammar theory

Language is necessary for the expression of rational thought and the organization of human societies. It may well have evolved from the need for social bonding between individuals belonging to the same group (Nadel, 1930). Language also permits projections in the past and in the future and is necessary for the transmission of knowledge (Leroi-Gourhan, 1988). While these characteristics among others, make language specific to *Homo sapiens,* they also seem to contribute to the splendid isolation of the linguistic function among the other human cognitive activities. Largely because of the enormous impact in cognitive sciences of the Generative Grammar Theory, developed by Chomsky (1957), language is often considered as relying on specific cognitive principles. Bickerton (2000), for instance, argues that the principles that govern language "seem to be specifically adapted for language and have little in common with general principles of thought or other apparatuses that might be attributable to the human mind" (p.158). Thus, one of the most important claims of the Generative Grammar (GG) theory is that language is autonomous from the other cognitive functions (Chomsky, 1957, 1991; Jackendoff, 1997; Pinker, 1994; Pollock, 1997). Language is considered as a computational module that entails its own functional and neural architecture (Molino, 2000). Moreover, the linguistic module comprises different sub-modules, each responsible for different levels of language processing reviewed above: phonology, morphology, syntax, semantics and pragmatics. Each sub-module is encapsulated (Fodor, 1983), such that the processing of information taking place within a module is performed independently of the processing of information in the other sub-modules. Phonological processing, for instance, is realized without being influenced by the morphological, syntactic, semantic or pragmatic aspects of language processing. Thus, the computations required to process language are specific to language, and the computations within one module are performed independently from those in the other modules.

Another basic claim of the GG theory is that languages are defined by their deep syntactic structure: syntax plays a dominant role in the structural organization of language. Moreover, from a functional perspective, syntax is first (Frazier, 1987). Logico-mathematic computations are first performed on symbols that have no intrinsic meaning; they only acquire meaning in a second step. Therefore, the chain of computations necessary to process language is considered as serially and hierarchically organized.

3.2. Other linguistic theories

It should be noted, however, that other linguistic theories have been developed in the last 20-30 years that advocate very different views of language structural and functional organization (François, 1998; Victorri, 1999). While it is beyond the scope of this chapter to go into the details of these different linguistic theories, that differ from each other in many respects (functional grammar: Dik, 1997; Van Valin and La Polla, 1997; cognitive grammar: Lakoff, 1987; Langacker, 1987; Talmy, 1988; linguistic functional typology; Croft, 1995; Greenberg, 1995; Givon, 1995), the important point is that these theories call into question the two basic claims of the GG theory summarized above. First, they reject the idea following which language is an autonomous function relying on its own structural and functional architecture. In contrast, they consider that language is an emergent property, relying on general cognitive principles, linked with perceptual and sensory-motor motor experiences (Fuchs, 1997; Robert, 1997). Second, they reject the syntactico-centrism of the GG and the idea of the autonomy of syntax relative to phonology, morphology, semantics and pragmatics (François, 1998). Following Langacker (1987), for instance, semantics, morphology, and syntax form a continuum with specific meaning associated to lexico-semantic units and schematic meaning associated to grammatical units. Thus, in contrast with the GG view that grammatical units are semantically empty, all grammatical elements have meaning. Moreover, linguistic units are not static but constructed through a dynamic process influenced by the context of enunciation (Culioli, 1999) and the interactions between individuals in a situation of communication (Fauconnier, 1997). Therefore, language should not be studied in isolation but rather in relation with other cognitive functions, specifically attention and short-term and episodic memory (Givon, 1995).

3.3. Success of Generative Grammar in Cognitive Neurosciences

Several reasons may explain the success of the GG theory both in linguistic and cognitive sciences, but two are of particular interest. First, the cognitive stakes of the GG theory have been clearly explained. It has therefore been possible to make predictions and design experiments to test these predictions (Chomsky, 1991; Robert, 1997). Second, the modular organization of the functional aspects of language processing is clearly neuro-compatible. The conception following which language is organized in sub-modules, each responsible for one specific processing stage, find strong support in the localisationist views of cerebral organization. The recent development of brain imaging methods, together with older data from the neuropsychological literature, largely contribute to the idea that specific functions are implemented in specific brain structures. However, a quick review of the literature will show that while this conception is probably correct regarding the mapping of basic sensory functions into the organization of primary, sensory brain areas, the story certainly becomes more complicated when trying to localize such higher order cognitive abilities as language or music.

4. Neuropsychology and brain imaging

Insofar as one agrees with the conception following which music and language cannot be considered as wholes but need to be subdivided into their component operations, it becomes unrealistic, for instance, to view the musical function as localized in the right hemisphere and language in the left. Rather, it may be that some aspects of music processing preferentially involve right cerebral structures, while others require structures on the left. The same remark applies to language as well. With this view in mind, the task of the cognitive neuroscientist is to delineate the different computations performed within one level of processing, to understand the mechanisms that underlie these computations and to localize where in the brain these mechanisms are implemented. Of course, this task is bound with both philosophical and methodological problems (Pacherie, 1999), but science is advancing rapidly and new methods are available to track these issues. We will now review some findings issued from studies in neuropsychology and in cognitive neurosciences relevant for our understanding of the functional and structural organization of music and language,.

4.1. The Neuropsychology of music

As the neuropsychology of music is a very recent field of research, the aim of most studies has often been very general. Thus, results appear still unsatisfying and fragmentary when compared to results obtained within other domains, such as language. At least four not exclusive factors may account for the specific difficulties encountered in this field. First, complex musical functions can only be studied in professional musicians, like complex linguistic functions can only be studied in humans. Unfortunately, music education in most countries is very poor, when compared to linguistic education. Thus, the study of impairment in music abilities has been limited by the scarcity of suitable patients. In saying this we are not arguing that the study of music deficits in non-musicians is uninteresting, but rather that it is limited to certain implicit musical abilities. Most importantly, when dealing with non-musicians, it is often difficult to know how musical was the patient premorbidly. Second,

most of the available case descriptions of amusia, while sometimes admirably detailed, lack a theoretical frame (Schön, Semenza and Denes, in press). Third, Basso (1999) clearly underlines the fact that "the encounter between a neuropsychologist and a brain-damaged musician has generally caused the study of the patient's musical ability, independently of the researcher's interest for music." (pp. 410). Thus, a brain-damaged musician rarely comes across a musically competent neuropsychologist. Fourth, no standardised test battery for music exists to our knowledge (exception made for the Seashore test, 1960, originally created to test musical talent, possibly available on long playing 78 rev/min!), while different standardised national and international neuropsychological tests exist within other domains (e.g. W.A.I.S, Raven progressive Matrices). Thus, it is unlikely that a hospitalised patient will be submitted even a general music test. In the lucky case he was, it would remain difficult to compare those results with results of other studies using completely different tests.

Keeping these remarks in mind, there are nevertheless interesting studies of amusic patients reported in the neuropsychological literature. The first neuropsychological studies of music contrasted global musical abilities and other abilities such as language. Several studies reported musical disturbances in absence of language deficits (Dorgeuille, cited in Peretz et al., 1997; Marin, 1982; Peretz, 1997; Basso, 1999) and vice-versa (Luria et al., 1965; Assal, 1973; Basso and Capitani, 1985; Signoret, 1987). This kind of double dissociations suggests both a functional independence and a structural independence, in the sense of underlying neural networks, of these two faculties. However, what is dissociated and what is independent from what, remain to be determined. As noted above, neither music nor language can be considered as single entities but both rather need to be decomposed into their component operations. For instance, recent results have demonstrated that, within the musical domain, specific subcomponents can be selectively damaged. Peretz and Kolinsky (1993), based on results obtained from a patient showing amelodia without arhythmia, argued for partial separability between melody and rhythm. In another study, Peretz (1990) advocated the idea that the left hemisphere is better equipped for dealing with local features of melody (intervals) and the right hemisphere for arriving at global melody representations (contour).

Other results have shown that patients with right temporal lobe lesions exhibited a significant deficit in musical timbre perception, in comparison to patients with left temporal lobe lesions and normal control subjects (Samson and Zatorre, 1994). Moreover, Peretz (1996) described a patient markedly impaired at naming a tune, judging its familiarity and memorising new musical materials. The author suggested the existence of a perceptual memory, specialised for music, that can be selectively damaged, therefore producing deficits in recognition ability. Results of a different patient showed that tonal interpretation of melodic sequences was impaired, while the ability to process temporal information, melodic contour and to some extent interval size was preserved (Peretz, 1993). The author interprets her findings as evidence for a modular organisation of tonal knowledge, and argues that tonal encoding “possesses several properties of a module in Fodor’s sense”.

Interestingly, there seems to be some confusion regarding the Fodorian concept of modularity. It is often considered that, to be a module, a cognitive system has to possess all the features proposed by Fodor. However, Fodor (1983) did not offer a definition of modularity (“I am not, in any strict sense, in the business of defining my terms”, pp. 37); nor was he proposing any criteria. Rather, as Coltheart (1999) points out, Fodor only suggested a number of properties that are typical of modular systems. Thus, an interpretation in terms of strong and complete modularity is misleading. Fodor himself speaks of degrees of modularity and refers to a modular cognitive system as meaning “to some interesting extent” (ib.).

4.1.1. Modularity of language and music processing?

As mentioned previously, the modularity of language processing is based on the idea that the computations necessary for analysing linguistic information differ from those

required within other cognitive domains. Indeed, within the information processing chain leading to meaning from a series of acoustic or visual signals, some computations are probably language specific. However, others are probably common to other non-linguistic auditory and visual signals, organised according to precise structural and functional rules. Nevertheless, the specificity of the processes involved within one domain is hardly ever tested in comparison with those involved in other domains. For example, studies on language impairments are rarely concerned with music impairments and vice-versa, and this holds for other researches as well. Therefore, this brings to the conclusion that, taking into account the current organisation of research, if interactions between domains were to exist, they would hardly be found.

Importantly, however, a few studies have been aimed at directly comparing music and language. While an exhaustive review is beyond the point we would like to make here, it is of interest to illustrate how such comparisons can be performed at different cognitive levels. To explore the relationship between the processing of melodic and rhythmic patterns in speech and music, Patel et al. (1998a) tested the prosodic and musical discrimination abilities of two "amusic" patients. Prosody is an important aspect of spoken language, that encompass the melodic variations of the voice (intonation), the accents and accentuations, as well as the rhythm and pauses within sentences (see **Cutler** et al, 1997 ; **Hirst** & Di Cristo, 1998, for reviews). This aspect has, however, long been neglected mainly due to the technical and methodological difficulties encountered in the analysis and manipulation of the speech signal. Prosodic discrimination was assessed using sentence pairs where members of a pair differed by intonation or rhythm. Musical discrimination was tested using musical-phrase pairs derived from the prosody of the sentence pairs. This materials was used in order to make task demands as comparable as possible across domains. Results showed that the level of performance was similar across domains, good on both linguistic and musical discrimination tasks for one patient, and poor for the other. These results suggest shared neural resources for prosody and music. One potential problem with this interpretation is linked with the lack of musicality of the musical phrases derived from the prosody of the sentence pairs. Such "musical phrases" may lack some musical structure. Moreover, rather than keeping the two domains separate in two different tasks, it could be more appropriate to design a task with both dimensions included. In doing so their interaction could then be tested. But this is by no mean an easy task!

A very good example of how this can be achieved is found in Peretz et al. (1997). Studying an amusic patient, the authors designed an experiment to directly compare language and music. They chose the song form, as words and music are tightly bound in this musical form. In a song, the linguistic and musical structures go together, at least from a metric point of view. Moreover, even though the words used in a song have already been heard in a non-musical context, lyrics and tunes have a comparable degree of familiarity, since they are mostly sung together. The authors argue that, since lyrics and tunes are necessarily associated in a song, it should be possible to find priming effects. Finding a priming effect only for words in the amusic patient, who could no longer recognise familiar music, would argue for the independence of language and music processing in songs. In contrast, the finding of a priming effect on both dimensions, even if smaller than for control subjects, would claim for an integrated representation of lyrics and tunes. Unfortunately, probably due to the fact that working with patients implies strong time constraints, they tested words priming on word targets and music priming on music targets, but they did not cross the two dimensions. Moreover, another problem arises when considering that music recognition is not as well temporally determined in time as word recognition. Even though this design still presents some problems, this task is a very good example of how music and language can be combined together to yield important information on how they are processed by the brain.

4.2. Evidence from brain imaging

Brain imaging methods are aimed at understanding the functional activity of the brain either directly through the measures of the electric activity of single neurons (intra-cellular recordings) or of neuronal populations (Electroencephalography, EEG and Magnetoencephalography, MEG by the analysis of the magnetic activity that is coupled with the electrical activity), or indirectly through the measures of the brain metabolic activity (Positon Emission Tomography, TEP, and functional Magnetic Resonance Imaging, fMRI). Overall, the direct methods have an excellent temporal resolution and a relatively poor spatial resolution, while the reverse is true for indirect methods. Very elegant works have been conducted using these different methods demonstrating, for instance, the retinotopic organization of the visual cortex using fMRI (Tootel, Hadjkhani, Mandola et al., 1998) and the tonotopic organization of the auditory cortex using intra-cellular recordings, MEG (Pantev, Hoke, Luetkenhoerner et al. 1989) or fMRI (Strainer, Ulmer, Yetkin et al., 1997). Thus, there is a strict mapping between the organization of the receptor fields at the periphery, either in the retina or in the cochlea, and the functional organization of the primary visual and auditory cortex. Aside from extending to humans previous discoveries in animals, these findings validate the use of such complex methods as fMRI to study human perception and cognition.

4.2.1. Semantic, melody and harmony

A starting point in the study of the neurophysiological basis of language processing has been the discovery of the N400 component by Kutas & Hillyard (1980). This negative component of the ERPs, peaking around 400 ms after word onset, is elicited by words that are semantically unexpected, incongruous, within a linguistic context (e.g., "The pizza was too hot to cry", see Figure 1). Further results have shown that N400 amplitude is modulated by semantic priming, so that an unexpected word related to the best sentence completion (e.g., “drink” when the expected word is “eat”, see Figure 1) elicit a smaller N400 than a completely unexpected word (e.g., “cry”; Kutas and Hylliard, 1984). These results, together with those issued from a large number of experiments, have lead to the concensus that the N400 is a good index of the integration process of a word within its linguistic context.

Please insert Figure 1 around here

The first experiments that we designed were aimed at finding out whether an N400 component would also be elicited when melodically and harmonically unexpected notes are presented within a melodic context (Besson and Macar, 1987; Besson , Faita and Requin, 1994; Besson and Faita, 1995). We presented both familiar and unfamiliar monodic musical phrases to musicians and non-musicians. The familiar melodies were chosen from the classical repertoire of Western occidental music from the 18th and 19th centuries, and the unfamiliar musical phrases were composed by a musician following the rules of tonal harmony (see Figure 2). These melodies were ended by the congruous or most expected note, by a note out of the tonality of the musical phrase (non-diatonic incongruities perceived as wrong notes) or by a note within the tonality but not the most expected ending (melodic or diatonic incongruities). Thus, we created a degree of musical incongruity from diatonic to non-diatonic

Please insert Figure 2 around here

Results clearly showed that both types of unexpected notes elicited the occurrence of late positive components, peaking around 600 ms (P600). As demonstrated for the N400 component, P600 amplitude was shown to depend upon the degree of musical incongruity: it was larger for the most unexpected, non-diatonic wrong notes than for the less unexpected, diatonic incongruities. Moreover, the amplitude of the P600 was larger for familiar than unfamiliar musical phrases and for musicians than for non-musicians (see Figure 3). These findings clearly demonstrate not only that specific notes are expected within a musical phrase, but that such expectations depend upon the familiarity of the musical excerpts and the expertise of the listener. Thus, one of the interesting similarities between language and music mentioned above, their ability to generate strong expectancies, is supported by empirical evidence. However, our results also show that the processes that govern semantic expectancy, and reflected by a negative component, peaking around 400 ms, the N400, are qualitatively different from those involved in musical expectancies, and reflected by a positive component peaking around P600 ms, the P600. While, to our knowledge, the functional significance of positive versus negative polarities in the ERPs is not clearly established, our results, by demonstrating qualitative differences between language and music processing, nevertheless strongly argue for the specificity of the processes involved in computing the semantic aspects of language. Thus, one of the most important differences between language and music outlined in the introduction, the fact that, in contrast to language, music has no intrinsic meaning and is a self-referential system, seems to find some support in these experimental findings.

Please insert Figure 3 around here

4.2.2. Semantic and harmony in opera

Opera is perhaps the most complete art form as it calls upon music, language, drama and choreography. It originated in Italy at the end of the 16th century with the “Euridice” of the Florentine composer Jacopo Corsi (1561-1602), played as a wedding gift to Maria de Medici and Henri IV (1600). Opera, as a new art form, then spread to other Italian courts with the better known “Orfeo” of Monteverdi (1604). Since this time, a question that interested both music analysts and composers has been to determine which of the words or the music plays the most important role in opera. In his "Life of Rossini", Stendhal (1783-1842) argued that music is most important: "its function is to animate the words". Later, ethnomusicologists, such as Levman (1992), have pointed out that the lyrics are subordinate to the music in tribal songs and rituals. In contrast, Richard Wagner (1813-1883) considered that both aspects are intrinsically linked: "Words give rise to the music and music develops and reinforces the language", an opinion shared by Boulez (1966): "The text is the center and the absence of the musical piece". Richard Strauss (1864-1949) once even composed an opera "Capriccio" (1940) to illustrate the complementarity of words and music.

To try to determine, based on scientific grounds, whether the words or the music play the most important role when we listen to opera, we selected two hundred excerpts from French operas from the 19th and 20th centuries (Besson, Faita, Peretz et al. 1998). Each excerpt lasted between 8 and 20 seconds and was sung a capella by a woman singer in each of four experimental conditions: the final word of the excerpt was semantically congruous and sung in tune, semantically incongruous and sung in tune, semantically congruous and sung out of tune and both semantically incongruous and sung out of tune (see Figure 4).

Please insert Figure 4 around here

Based on previous results (Kutas and Hillyard, 1980), it was of interest to determine whether semantically incongruous words will also elicit an N400 component when they are sung. Similarly, it was of interest to determine wether congruous words sung out of tune will also elicit a P600 component (Besson and Faita, 1995). Of most interest, was the double incongruity condition: will semantically incongruous words sung out of key elicit both an N400 and a P600 component? If language plays the most important role when we listen to opera, then results may show an N400 but no P600. Conversely, if music is the cornerstone of opera, then results may show a P600 without N400. But may be both effects will be elicited; they may then be additive (i.e., equal to the sum of the effect associated with each type of incongruity alone) or interactive. To answer these questions, we recorded the ERPs associated with the final words of each excerpt, from 16 professional musicians from the opera in Marseille.

To summarize, results demonstrated that sung incongruous words did elicit an N400 component, thus extending to songs results previously reported for written and spoken language (Kutas and Hillyard, 1980; MacCallum, Farmer and Pocock, 1984); see Figure 5A. Moreover, words sung out of tune did elicit a P600 component, thus extending to songs results previously reported for out of tune notes (Besson and Faita, 1995; Paller, McCarthy and Wood, 1992; Regnault, Bigand and Besson, 2000); see Figure 5B. Most interesting are the results in the double incongruity condition. They show that incongruous words sung out of tune elicit both a N400 and a P600 components (see Figure 5C). Interestingly, the N400 occurred earlier than the P600, which is taken as evidence that the words were processed faster than the music. Finally, effects in the double incongruity condition were not significantly different from the sum of the effects observed in each condition of simple incongruity (see Figure 6). This finding provide a strong argument in favor of the independence (i.e., the additivity) of the computations involved in processing the semantic aspects of language and the harmonic aspects of music. Therefore, when we listen to opera, we process both the lyrics and the tunes in an independent fashion, and language seems to be processed before music.

Please insert Figures 5 and 6 around here

4.2.3. The influence of attention

We tracked these results further by conducting another series of experiments aimed at studying the effect of attention, again testing some professional musicians from the opera in Marseille.[76] We hypothesized that if lyrics and tunes are processed independently, listeners should be able to focus their attention only on the lyrics or only on the tunes depending upon the instructions. Without going into the details of the results, an N400 component was elicited to sung incongruous words and a P600 was associated with congruous words sung out of tune, thus replicating our previous findings (Besson, Faita, Peretz et al. 1998). Most interestingly, the N400 to incongruous words completely vanished when participants focused their attention on the music (see Figure 7). Thus, musicians were able not to process the meaning of words; they did not notice whether or not the terminal word made sense within the linguistic context when they only listened to the music. Conversely, P600 amplitude was significantly reduced when musicians focused attention on language, so that they did not hear that the final word was sung out of tune. Taken together these results again provide strong arguments in favor of the independence of lyrics and tunes. There is some limit to such processing independence, however. Results in the double incongruity condition showed that the presence of one type of incongruity influenced the processing of the other type of incongruity. When words were both semantically incongruous and sung out of tune, musicians could not help but hearing the musical incongruity even if they were asked to focus their attention on language.

Please insert Figure 7 around here

4.2.4. Syntax and harmony

The rules of harmony and counterpoint are often described as the grammar of tonal music. As syntax is used to extract the fundamental structure of an utterance by assigning different functions to different words, the rules of harmony allow to specify the different elements, notes and chords, that fulfill a specific harmonic function. Results of experiments manipulating the harmonic function of target chords have shown that violations of harmonic expectancies are associated with P600 components (Janata, 1995; Regnault, Bigand and Besson, 2000). Interestingly, research on syntax using ERPs has also shown that different types of syntactic violations, such as violations of gender, word order or noun-verb agreement, elicit a positive component, peaking around 600 ms (Hagoort, Brown and Groothusen, 1993; Osterhout and Holcomb, 1992; Friederici 1998). Moreover, both components show a similar parietal distribution over the scalp, which, together with their similar polarity and latency, seems to indicate that they reflect qualitatively similar processes.

In order to further test this hypothesis, Patel & collaborators (Patel, Gibson and Ratner, 1998) conducted an experiment directly aimed at comparing the P600 components elicited by harmonic and syntactic (phrase structure) violations. ERPs associated to a word within a grammatically simple, complex or incorrect sentence were compared to those associated with the presentation of a chord that belonged to the same, a nearby or a distant tonality than the one induced by the chords sequence. Results showed that, aside from early morphological differences in the ERPs to words and chords, due to the differences in the acoustic characteristics of these two types of auditory signals, the effects associated with the violation of syntactic and harmonic expectancies were not significantly different (see Figure 8). Therefore, these results raise the interesting possibility that a general cognitive process is called into play when participants are asked to process the structural aspects of an organized sequence of sounds, be it language or music. Finally, an early right anterior negativity was found around 300-400 ms in response to a chord belonging to a distant tonality. These results parallelled those obtained in language experiments showing that an early left anterior negativity is also associated with some syntactic violations (Friederici, Pfeifer and Hahne, 1993). While these two negative components showed a different distribution over the scalp, with a left predominance for language and a right predominance for music, they may reflect functionally similar processes.

Please insert Figure 8 around here

4.2.5. Temporal structure

Spoken language, as music, is composed of acoustic events that unfold in time. Because of the temporal structure inherent to both language and music, specific events are expected at specific times. The main question addressed in the next series of experiments was to determine whether the processes involved to analyze temporal structures rely on general cognitive mechanisms, or rather differ as a function of the specific characteristics of the materials to be processed. We used both the ERP and the MEG methods to analyze the time course of the effects of temporal structure violations in language and music, and fMRI to localize the cerebral structures activated by these violations. We hypothesized that, if a general mechanism is responsible for processing of the temporal structures in language and

music, qualitatively similar effects should be revealed in the ERP and MEG recordings, and similar brain areas should be shown to be activated, by temporal violations. In contrast, if processing temporal information in both systems rely on different mechanisms, qualitatively different effects, and different brain areas should be found in language and music.

In previous experiments (Besson and Faita, 1995) we introduced an unexpected silence between the before to the last and the last note of a musical phrase (see Figure 2). Results showed that a large biphasic, negative then positive, potential, the emitted potential (Sutton, Braren, Zubin et al., 1967), was elicited when the final note should have been presented but was not since it was delayed by 600 ms. The amplitude of this effect was similar for musicians and non-musicians, but was larger for familiar than unfamiliar melodies (see Figure 9). These findings clearly indicate that both musicians and non-musicians were able to anticipate the precise moment when the final note was to be presented, and were surprised when it was not. Moreover, knowing the melodies allowed participants to generate more precise expectancies than when melodies were unfamiliar. Therefore, these results indicate that the occurrence of an emitted potential can serve as a good index of temporal expectancy.

Please insert Figure 9 around here

It was then of interest to determine whether similar results would be found for spoken language (Besson, Faita, Czernasty et al., 1997). To this aim, we presented both familiar (e.g., proverbs) and unfamiliar auditory sentences to participants. In half of sentences, final words occurred at their normal position, while in the other half, they were delayed by 600 ms. Results showed that an emitted potential, very similar to the one described for temporal ruptures in music, developed when the final word should have been presented (see Figure 10).

Therefore, these ERP results indicate that qualitatively similar processes seem to be responsible for temporal processing in language and music.

Please insert Figure 10 around here

In order to strengthen this interpretation, it was important to determine whether the same brain structures are activated by the processing of temporal ruptures in language and music. As mentioned above, fMRI allows to localize brain activation with an excellent spatial resolution. Moreover, the MEG permits to localize the generators of the effects observed on the scalp more precisely than the ERP method, while also offering an excellent temporal resolution. Therefore, in collaboration with Prof. H. Heinze and his research team, we conducted three experiments, in which we presented both auditory sentences and musical phrases (Weyert, Besson, Tempelmann et al.). These experiments used a blocked design in which only sentences or musical phrases without temporal ruptures were presented within a block of trials, and only sentences or musical phrases with temporal ruptures at unpredictable positions were presented within another block of trials. The ERP method was used in the first experiment to replicate, within subjects, the results found previously with two different groups of subjects (Besson and Faita, 1995; Besson, Faita, Czernasty et al., 1997), and the fMRI and the MEG methods were used respectively in the other two experiments, trying to localize the effects of interest.

Overall, the ERP results replicated, within subjects, those previously found in music and language separately (i.e., an emitted potential). However, the comparison of the conditions with and without temporal violations revealed a somewhat different pattern of activation using the MEG and fMRI methods. Source localization based on the MEG data revealed that the underlying generators of the biphasic potential recorded on the scalp were

most likely located in the primary auditory cortex of both hemispheres. In contrast, fMRI results showed activation of the associative auditory cortex in both hemispheres, as well as some parietal activation. Several factors may account for these differences, but the main point is that similar brain areas were activated by temporal violations in both language and music. Therefore, taken together our results suggest that processing temporal information in both language and music rely on general cognitive mechanisms.

5. Conclusion

In this chapter we have addressed one of the central question of human cognition, the specificity of language processing. Is language an autonomous system, independent from other human cognitive abilities or does language rely on general cognitive principles? To address this question, we have conducted several experiments aimed at comparing some aspects of language processing with some aspects of music processing. We mainly used the Event-Related Potentials method, that offers an excellent temporal resolution, and therefore permits to study the time course of information processing, and to determine whether the processes involved in language and music are qualitatively similar or different.

Taken together, results have shown that the semantic computations required to access the meaning of words, and their integration within a linguistic context, seem to be specific to language. Indeed, while unexpected words within a sentence context are associated with the occurrence of an N400 component, unexpected notes or chords within musical phrases elicit a P600 component. In contrast, words that are unexpected on the basis of the syntactic structure of the sentence, and chords that are unexpected as a function of the harmonic structure of the musical sequence, elicit similar effects in both cases, namely a P600 component. Early negative effects, the Left Anterior Negativity and the Right Anterior Negativity, that developed between 200-300 ms, have also been reported in experiments manipulating syntax and harmony, respectively. While their different scalp distribution seems to indicate that they reflect the involvement of different brain structures, more research is needed to further track their functional significance. Finally, violations of temporal structure within language and music also elicit similar effects, a biphasic negative-positive complex, the emitted potential. The occurrence of the emitted potential shows that, in both language and music, words and notes or chords are expected at specific moments in time. Therefore, when we listen to language and music we do not only expect words or chords, with specific meaning and function, but we also expect them to be presented on time!

The question of the specificity of language processing has broad implications for our understanding of the human cognitive architecture, and even more generally, for the fundamental problem of the relationship between structures (the different brain regions) and functions (language, music, …). While researches reported here shed some light on some aspects of language processing and highlight some of the similarities and differences with music processing, more research clearly needs to be conducted within this fascinating research domain. It is of most interest to combine brain imaging methods that offer an excellent temporal resolution, as the ERPs method used in the experiments described above and TEP and fMRI that offer and excellent spatial resolution, in order to pinpoint the spatio-temporal dynamics of the networks of cerebral structures involved when we are engaged in two of the most human cognitive abilities: language and music.

Some experiments have already been designed to directly compare language and music using fMRI and TEP. Binder et al. (Binder, Frost, Hammeke et al., 1996) compared tones and words processing in an fMRI study. Results showed that several brain structures, including the left superior temporal sulcus, the middle temporal gyrus, the angular gyrus and the lateral frontal lobe, showed stronger activation for words than tones. However, both types of stimuli activated the Heschl gyrus and the superior temporal plane, including the planum temporale (PT). The authors concluded that while the PT is similarly involved in the auditory processing of words and tones, other broadly distributed areas are specifically involved in word processing. Gandour et al. (Gandour, Wong and Hutchins, 1998) conducted a PET study in which both Thai and English participants were required to discriminate pitch patterns and Thai lexical tones derived from accurately filtered Thai words. Results of the tone minus pitch subtraction indicated that only native Thai speakers showed activation in the left frontal operculum (BA44/45). This finding was taken as evidence that Thai lexical tones are meaningful for native Thai speakers, but not for English speakers. However, for our purposes, it is also interesting to note that for both Thai and English speakers, several structures, the left anterior cingulated gyrus (BA32), the left and right superior temporal gyrus (BA 22) and the right cerebellum, were activated both in the pitch and in the tone tasks.

More generally, results have shown that primary auditory regions (BA 41 & 42) responds in similar ways to speech and music (Zatorre, Evand, Meyer et al., 1992). Secondary auditory regions (BA22) are activated by hearing and understanding words (Falk, 2000) as well as by listening to scales (Sergent, Zuck, Terriah et al., 1992), auditory imagery for sounds (Zatorre, Halpern, Perry et al., 1996) and access to melodic representations (Platel, Price, Wise et al, 1997). The supramarginal gyrus (BA 40) seems involved in understanding the symbolism of language (Falk, 2000) and the reading of musical scores (Sergent et al., 1992). Broca’s area is known to be involved in motor activity related to language, and was also shown to be active when playing music (Sergent et al., 1992) and when musicians were engaged in a rhythmic task (Platel et al., 1997). The supplementary motor areas (BA 6) and the right cerebellum are also active when playing and imaging playing music (Sergent et al., 1992; Chen, Kato, Zhu et al., 1996). While this list is far from exhaustive, it nevertheless suffices to show that some of the most important language areas are clearly involved in music processing as well. Of course, some brain structures also seem to be specifically or preferentially involved in language processing (Grabowsky and Damasio, 2000) and the converse is true for music (Zatorre and Binder, 2000).

However, in order to have a better understanding of the specific brain structures involved in language processing, one would need a meta-analysis of the results obtained across the many experiments aimed at localizing the different aspects of language processing. One would then need to do the same for music, or for any other cognitive function of interest, and then compare the results of these meta-analyses. While such meta-analyses are starting to be performed for some aspects of language processing, such as language production (Indefrey and Levelt, 2000) or prelexical and lexical processes in language comprehension (Norris and Wise, 2000), the data in the neuroimaging of music are still too scarce for such an enterprise. Moreover, assuming that such meta-analyses are performed for music as well, it will still remain extremely difficult to compare results of experiments that were not directly designed to compare language and music processing. Indeed, even leaving aside the theoretical problem of which level of processing in language is best compared with which level of processing in music, the choice of the task to be performed on the stimuli, its difficulty, as well as experimental factors, such as the mode (blocked versus mixed) and rate of stimulus presentation, stimulus repetition, and data analysis (subtraction method, correlative analyses, etc…), have been shown to exert a predominant influence on the results obtained. As progress in neuroimaging is extremely rapid, one can hope that these difficulties will be overcome in a very near future.

References

Aiello, R. 1994. Music and language: parallels and contrasts. *In* Musical perceptions. Aiello R. & J. Sloboda, Eds.: 40-63. Oxford University Press. New York, NY.

Assal G, Buttet J. 1983Agraphie et conservation de l'écriture musicale chez un professeur de piano bilingue. Rev. Neurologique, 139, 10: 569-574.

Basso A, Capitani E. 1985Spared musical abilities in a conductor with global aphasia and ideomotor apraxia. J-Neurol-Neurosur-Psychiatry, 48: 407-412.

Basso A. 1999. The Neuropsychology of Music. In: Denes G, Pizzamiglio L, editors. Handbook of Neuropsychology, Hove: Psychology Press.

Besson, M & F. Faïta. 1995. An Event-Related Potential (ERP) study of musical expectancy: Comparison of musicians with non-musicians. Journal of Experimental Psychology: Human Perception and Performance 21, 6: 1278-1296.

Besson, M, F. Faïta.& J. Requin. 1994. Brain waves associated with musical incongruity differ for musicians and non-musicians. Neuroscience Letters 168: 101-105.

Besson, M. Macar, F. 1987. An event-related potential analysis of incongruity in music and other non-linguistic contexts. Psychophysiology 24, 1: 14-25.

Besson, M., F. Faïta, C. Czternasty & M. Kutas. 1997. What's in a pause: event-related potential analysis of temporal disruptions in written and spoken sentences. Biological Psychology 46: 3-23.

Besson, M., F. Faita, I. Peretz, A. M. Bonnel et al. 1998. Singing in the brain: Independence of lyrics and tunes. Psychological Science 9, 6: 494-498.

Bickerton, D. 2000. Can Biomusicology Learn from Language Evolution Studies? *In* The Origins of Music. N.L. Wallin, B. Merker & S. Brown, Eds.: 153-164. MIT Press. Cambridge, MA.

Binder, J. R, J. A. Frost, T.A. Hammeke et al. 1996. Function of the left planum temporale in auditory and linguistic processing. Brain 119: 1239-1247.

Blacking, J. 1973. How Musical is Man? University of Washington Press. Seattle.

Boucourechliev, A. 1993. Le langage musical. Collections les chemins de la musique. Fayard, Paris

Boulez, P. 1966. Relevés d'apprenti. Editions du Seuil, Paris.

Chen, W., T. Kato, X.H. Zhu, et al. 1996. Functionnal mapping of human brain during music imagery processing. NeuroImage, 3, S205.

Chomsky, N. 1957. Syntactic structures. Mouton & Co. The Hague, NL.

Chomsky, N. 1988. Language and problems of knowledge. The Managua lectures. MIT Press. Cambridge, MA.

Chomsky, N. 1991. Linguistics and cognitive science: Problems and mysteries. *In* The Chomskyan Turn. A. Kasher, Ed.: 26-53. Basil Blackwell. Cambridge, MA.

Coltheart, M 1999. Modularity and cognition. Trends in Cognitive Neuroscience, 3, 3, 115-120.

Croft, W. 1995. Autonomy and functionalist linguistics. Language 71: 490-532.

Culioli , A.. 1999. Pour une linguistique de l'énonciation. Ophrys, Paris.

Darwin, C. 1871. The Descent of Man, and Selection in Relation to Sex. Murray. London.

Dik, S. 1997. The theory of Functional Grammar. Mouton De Gruyter, Berlin.

Eco, U. 1989. The Open Work. Cambridge, MA: Harvard University Press.

Eco, U. 1979. Trattato di semiotica generale. Bompiani. Milano.

Falk, D. 2000. Hominid brain evolution and the origin of music. *In* The Origins of Music. N.L. Wallin, B. Merker & S. Brown, Eds.: 197-216. MIT Press. Cambridge, MA.

Fauconnier, G. 1997. Mappings in Thought and Language, Cambridge University Press.

Fodor, J. 1983. Modularity of mind. MIT Press, Cambridge, MA.

François, J. 1998. Grammaire fonctionnelle et dynamique des langues - de nouveaux modèles d'inspiration cognitive et biologique. Verbum XX, 3, 233-256.

Frazier, L.. 1987. Sentence processing: A tutorial review. *In* Attention and performance XII. M. Coltheart Ed.: 559-586. Erlbaum. Hillsdale, NJ.

Friederici, A. D. 1998. The neurobiology of language comprehension. *In* Language comprehension: A biological approach. A.D. Friederici, Ed..: 263-301. Springer, New York.

Friederici, A.D., E. Pfeifer & A. Hahne. 1993. Event-Related brain potentials during natural speech processing: Effects of semantic, morphological and syntactic violations. Cognitive Brain Research 1: 182-192.

Fuchs, C. 1997. Diversité des représentations linguistiques : Quels enjeux pour la cognition ? *In* Diversité des langues et représentations cognitives C. Fuchs & Robert S., Eds. : 5-24. Ophrys.

Gandour, J., D. Wong & G. Hutchins 1998. Pitch processing in the human brain is influenced by language experience. Neuroreport 9: 2215-2119.

Givón, T. 1995. Functionalism and Grammar. Benjamins, Amsterdam.

Grabowski, T.J., A.R. Damasio 2000. Investigating language with functional neuroimaging. In Brain Mapping: The systems. A. W. Toga & J. C. Mazziotta, Eds.: 425-458. Academic Press.

Greenberg,, J. 1995. The diachronic typological approach to language. *In* Approaches to language typology. M. Shibatani &. Bynon T., Eds.: 145-166. Clarenton, Oxford.

Hagoort, P., C. Brown, & J. Groothusen. 1993. The syntactic positive shift as an ERP-measure of syntactic processing. Language and Cognition Processes 8: 439-483.

Halpern, A. R. and G. H. Bower 1982. Musical expertise and melodic structure in memory for musical notation. American-Journal-of-Psychology 95, 1: 31-50.

Hornbostel E.M. von (1928) African negro music, Africa: 1 (1), 30-62.

Indefrey, P. & W. J. M Levelt. 2000. The neural correlates of language production. *In* The new cognitive Neurosciences. M.S. Gazzaniga, Ed.: 845-865. MIT Press, Cambridge, MA.

Jackendoff R. 1997. The Architecture of the Language faculty, MIT Press, Cambridge, MA.

Jackendoff, R. 1991. Musical Parsing and Musical Affect. Music Percetion 9 (2): 199-230.

Jakobson, R. 1973. Essais de linguistique générale. II. Rapports internes et externes du langage. Editions de Minuit, Arguments. Paris.

Janata, P. 1995. ERP measures assay the degree of expectancy violation of harmonic contexts in music. Journal-of-Cognitive-Neuroscience 7, 2: 153-164.

Kivy, P. 1991. Music Alone. Philosophical Reflection on the Purely Musical Experience. Cornell Paperbacks.

Kubik, G. 1969. Composition techniques in Kiganda xylophone music. The African music journal 4 (3).

Kutas, M & Hillyard, S.A. 1984. Event-related brain potentials (ERPs) elicited by novel stimuli during sentence processing. Annals-of-the-New-York-Academy-of-Sciences 425: 236-241.

Kutas, M & S. A. Hillyard. 1980. Reading sens..ess sentences: Brain potentials reflect semantic incongruity. Science, 207, 203-205.

Lakoff, G. 1987. Women, Fire and Dangerous Things, University of Chicago Press.

Langacker, R. W. 1987. Foundations of Cognitive Grammar. Vol. I: Theoretical Prerequisites. Stanford University Press.

Leipp, E. 1977. La machine à écouter: Essais de psychoacoustique. Masson, Paris.

Leroi-Gourhan, A. 1988. Le geste et la parole. Vol. I: La mémoire et les rythmes. Sciences d'aujourd'hui, Albin-Michel, Paris.

Levman, B. G. 1992. The genesis of music and language. Ethnomusicology 36, 2: 147-170.

Locke S and Kellar L. Categorical perception in a non-linguistic mode. 1973. Cortex, 9, 35-369.

Luria A, Tsvekova L and Futer J 1965. Aphasia in a composer. Journal of Neurological Science, 2: 288-292.

Signoret JL, Van Eeckhout P, Poncet M, Castaigne P. 1987. Aphasie sans amusie chez un organiste aveugle. Rev. Neurologique, 143, 3: 172-181.

MacCallum, W. C., S. F. Farmer, P. V. Pocock. 1984. The effects of physical and semantic incongruities on auditory event-related potentials. Electroencephalography and Clinical Neurophysiology 59: 477-488.

Marin O. 1982. Neurological aspects of music perception and performance. In Deutsch D, editor. The Psychology of music. New York: Academic Press.

Meyer, L. 1956. Emotion and meaning in music. University of Chicago Press.

Molino, J. 2000. Toward an Evolutionary Theory of Music and Language. *In* The Origins of Music. N.L. Wallin, B. Merker & S. Brown, Eds.: 165-176. MIT Press. Cambridge, MA.

Nadel, S. 1930. The Origin of Music. Musical Quarterly 16, 531-546.

Norris, D. & R. Wise. 2000. The study of prelexical and lexical processes in comprehension: Psycholinguistics and functional neuroimaging. *In* The new cognitive Neurosciences. M. S. Gazzaniga, Ed.: 867-880. MIT Press. Cambridge, MA.

Osterhout, L. & P. J. Holcomb. 1992. Event-Related brain potentials elicited by syntactic anomaly. Journal of Memory and Language 31: 785-804.

Pacherie, E. 1999. Philosophie et Sciences Cognitives. *In* Encyclopédie Philosophique Universelle, J. F. Mattéi, Ed. PUF, Paris.

Paller, K. A., G. McCarthy, C. C. Wood.1992. Event-related potentials elicited by deviant endings to melodies. Psychophysiology 29 2: 202-206.

Pantev, C., M. Hoke, B. Luetkenhoener, et al. 1989. Tonotopic organization of the auditory cortex: Pitch versus frequency representation. Science 246, 4929: 486-488.

Patel A, Gibson E, Ratner J, Besson M, Holcomb P 1998. Processing syntactic relations in language and music: An event-related potential study. Journal-of-Cognitive-Neuroscience, 10 (6): 717-733.

Patel A, Peretz I, Tramo M, Labreque R 1998a. Processing prosodic and musical patterns: A neuropsychological investigation. Brain-and-Language., 61 (1): 123-144.

Patel, A. D., E. Gibson, J. Ratner. 1998. Processing syntactic relations in language and music: An event-related potential study. Journal of Cognitive Neuroscience 10, 6: 717-733.

Peretz I. Processing of local and global musical information by unilateral brain-damaged patients. Brain, 113: 1185-1205, 1990.

Peretz I. Auditory agnosia: a functional analysis. In: Mc Adams S, Bigand E, editors. Thinking in sound: the cognitive psychology of human audition. Oxford University Press, 1993, 199-230.

Peretz I. 1996. Can we lose memory for music? A Case of music agnosia in a nonmusician. Journal of Cognitive Neuroscience, 8, 6: 481-496.

Peretz I, Morais J 1989. Music and Modularity. Contemporary Music Review, 4: 279-293.

Peretz I, Kolinsky R. 1993. Boundaries of Separability between Melody and Rhythm in Music Discrimination: A Neuropsychological Perspective. The Quarterly Journal of Experimental Psycholog,; 46A, (2): 301-325.

Peretz I, Belleville S, Fontaine S. 1997. Dissociations entre la musique et le langage après atteinte cérébrale: un nouveau cas d'amusie sans aphasie. Revue Canadienne de psychologie expérimentale, 51, 4: 354-367.

Piana, G. 1991. Filosofia della musica. Guerini e associati, Milano.

Pinker, S. 1994. The language instinct: How the mind creates language. Harper Perennial.

Platel, H., C. Price, J. C. Wise et al. 1997. The structural components of music perception. Brain 120, 229-243.

Pollock, J. Y. 1997. Langage et cognition. Introduction au programme minimaliste de la grammaire générative. PUF, Paris.

Regnault, P., E. Bigand, M. Besson. 2000. :in press. Different brain mechanisms mediate sensitivity to sensory consonance and harmonic context: Evidence from auditory Event Related brain Potentials. Journal of Cognitive Neuroscience.

Robert, S. 1997. Variation des représentations linguistiques : des unités à l'énoncé. *In* Diversité des langues et représentations cognitives C. Fuchs & Robert S., Eds. : 25-39. Ophrys, Paris.

Samson S., Zatorre R. J 1994. Contribution of the right temporal lobe to musical timbre discrimination. Neuropsychologia 32, 2: 231-240.

Schön, A. & D. Schön. 1999. Il potere del suono e della musica. Fuga a più voci. Psiche (1-2): 159-165.

Schön, D., C. Semenza, & G. Denes 2001 Naming of musical notes:A selective deficit in one musical clef. Cortex, in press.

Sergent, J., E. Zuck, S. Terriah et al.1992.Distributed Neural Network Undelying Musical Sight-Reading and Keyboard Performance. Science 257, 106-109.

Sloboda, J. A. 1985. Musical Mind. The Cognitive Psychology of Music. Oxford: Oxford University Press.

Sloboda, J. A. 2000. Individual differences in music performance. Trends in Cognitive Sciences 4,10: 397-403.

Spencer, H. 1857. The origin and function of music. Fraser’s Magazine 56: 396-408.

Strainer, J. C., J. L. Ulmer, F. Z. Yetkin et al. 1997. Functional magnetic resonance imaging of the primary auditory cortex: Analysis of pure tone activation and tione discrimination. Am. J. Neuroradiol 18: 601-610.

Sutton, S., M. Braren, J. Zubin & E. R. John. 1967. Evoked potential correlates of stimulus uncertainty. Science 150: 1187-1188.

Talmy, L. 1988. The relation of grammar to cognition, *In* B. Rudzka-Ostyn, Ed. Topics in Cognitive Linguistics, Benjamins.

Tootell, R .B. H., N. K., Hadjkhani, J. D. Mandola et al. 1998. From retinotopy to recognition: fMRI in human visual cortex. Trends in Cognitive Sciences 2, 5: 174-183.

Van Valin R. D. & R. J. LaPolla 1997. Syntax Structure, meaning and function. Cambridge University Press.

Victorri, B. 1999. Le sens grammatical. In Languages 136 : 85-105.

Weyert, H., M. Besson, C. Tempelmann, M. Scholz, et al. An analysis of temporal structure in language and music using ERPs, MEG and fMRI techniques: in preparation.

Wittgenstein, L. 1953. Phylosophical investigations. Ed G E M Anscombe and R Rhees.

Zatorre, J. & J. R. Binder. 2000. Functional and Structural Imaging of the Human Auditory System. In Brain Mapping: The systems. A.W. Toga & J. C. Mazziotta, Eds.: 365-402. Academic Press.

Zatorre, R. A. C. Evans, E. Meyer et al. 1992. Lateralization of Phonetic Pitch Discrimination in Speech Processing. Science 256, 846-849.

Zatorre,R. J.,; A. Halpern, D. W. Perry et al. 1996. Hearing in the mind's ear: A PET investigation of musical imagery and perception. Journal-of-Cognitive-Neuroscience l 8, 1: 29-46.

Figure Legends

Figure 1: ERPs elicited by sentence-final words at the central recording site (Cz) for congruous and incongruous words and for incongruous words that are semantically related to the best sentence completion. The amplitude of the negative component, peaking at 400 ms post-final word onset (N400) is largest for incongruous words, intermediate for incongruous words related to the best sentence completion and smallest for congruous words. In this and subsequent figures, amplitude (µV) is represented on the ordinate, with negative voltage up, and time (ms) on the abscissa *(Adapted from Kutas & Hillyard, 1984).*

Figure 2: Examples of the stimuli used in the experiment *(Adapted from Besson & Faïta, 1995).*

Figure 3: ERP results for musicians and non-musicians are presented separately for familiar and unfamiliar musical phrases. The vertical lines mark the onset of the final note. Results are from one typical recording site, the parietal location (Pz). The amplitude of the positive component, P600, is larger for nondiatonic than for diatonic incongruity, for musicians than for non-musicians and for familiar than for unfamiliar musical phrases *(Adapted from Besson & Faïta, 1995).*

Figure 4: Example of the opera's excerpts used in the experiment. Approximate translation of the excerpts, from “Les Huguenots” (Meyerber): “Really, his naïvity is charming. However, he trembles in front of beautiful eyes”, and from "Faust" (Gounod) : "For me, the pleasures and young mistresses, the crazy orgy of the heart and the senses". Note that, in french, the final incongruous words “boeufs” and "sciences" rhyme with the expected

completions ‘yeux” and "sens". The final note of the excerpt is in or out of tune *(Adapted from Besson, Faïta, Peretz, Bonnel & Requin, 1998).*

Figure 5: ERPs results averaged across 16 professional musicians and recorded from the parietal electrode (Pz). Terminal congruous words sung in key (Cong./Cong.) are compared to (A) semantically incongruous words sung in tune (Incong./Cong.), (B) semantically congruous words sung out of tune (Cong./Incong.)and (C) semantically incongruous words sung out of tune (Incong./Incong.). The vertical lines mark the onset of the final word of the excerpts. A large N400 component develops in the 50 - 600 ms that follow the presentation of semantically incongruous words (A). In marked contrast, a P600 develops in the 400 - 1200 ms that follow the presentation of words sung out of tune (B). Most importantly, both an N400 and a P600 develop in response to the double incongruity (C; *Adapted from Besson, Faïta, Peretz, Bonnel & Requin, 1998).*

Figure 6: Overlapped are the ERPs to congruent and incongruent endings, recorded at the central recording site (Cz), when participants payed attention only to the language (left column) or only to the music (right column) of the opera's excerpts. A large N400 effect is generated when participants focus their attention on language. This effect completely vanish when attention is focussed on music (top row). Similarly, the P600 effect is much larger when participants payed attention to music than when they payed attention to language (medium row). Finally, when words are both semantically incongruous and sung out of tune, the N400 effect is larger when participants payed attention to the language and the P600 effect is larger when they payed attention to the music (bottom row; *from Regnault & Besson, in preparation).*

Figure 7: Left side: Examples of the sentences presented in the auditory language experiment. Results showed an increased positivity from the simple to the ungrammatical sentences. Right side: Representation of the circle of fiths. Examples of the stimuli used in the experiment. The target chord, shown by the downward-pointing vertical arrow, is the congruous chord. The two arrows below the musical notation point to moderately incongruous (nearby key) and highly incongruous (distant key) target chords. Results also showed an increased positivity from the in key chords to the distant-key chords *(Adapted from Patel, Gibson, Ratner, Besson & Holcomb, 1998).*

Figure 8: Overlapped are the ERPs to congruous notes and to the rhythmic incongruities ending familiar and unfamiliar musical phrases for musicians and non-musicians. Recordings are from the parietal electrode (Pz). Large emitted potentials are elicited when the final note should have been presented (vertical bar) but was delayed by 600 ms. The arrow points to the moment in time when the final note was presented *(Adapted from Besson & Faïta, 1995).*

Figure 9: Comparison of the effects of temporal violations in language and music. Recordings are from the parietal electrode (Pz). Left side: Overlapped are the ERPs to congruous words and to the temporal disruptions ending familiar and unfamiliar sentences. In both language and music, large emitted potentials are elicited when the final event should have been presented (vertical bar) but was delayed by 600 ms. Note that the amplitude of the emitted potential is larger in music than in language, but that in both cases, its amplitude is larger for familiar than unfamiliar materials (*Adapted from Besson, Faita, Czternasty & Kutas, 1997 and Besson & Faïta, 1995)*