Music and language: theories of common origins and related experimental data

Bruno T. Ruviaro

ruviaro@dartmouth.edu
Dartmouth College/Electro-Acoustic Music
6242 Hallgarten Hall
Hanover, NH 03755, USA

Abstract

This paper briefly summarizes basic considerations on evolutionary theories of music (Huron, 2001; Cross, 2003) as well as describes the musilanguage model — a theory that suggests a common ancestor for music and language (Brown, 2001). Further comparisons between the two domains are examined under a cognitive perspective, and recent experimental data on the inter-relation of music and language (more specifically, speech) are reviewed.

Keywords: musilanguage; evolutionary musicology; speech and music;

1. Introduction

There is a significant amount of recent research attempting to move forward the quest about the relationship between music and language. Studies range from general theories and models describing their possible common origins to specific experimental approaches trying to elucidate the details of such relationship. Although it is hard for someone to object to the striking similarities of these two highly specialized human-specific manifestations (music and language), it is still an open question as to whether these activities share common cognitive structures. In a review on the comparison of music and language, Besson and Schön (2001) list different theories that claim — in contrast with the Generative Grammar Theory — that language may actually be dependent on other general cognitive functions.

We start this paper by describing the musilanguage model (a theory that suggests a common ancestor for music and language — see Brown, 2001), as well as summarizing basic considerations of evolutionary theories of music (Huron, 2001; Cross, 2003). Further comparisons between the two domains are examined under a cognitive perspective, and recent experimental data on the inter-relation of music and language (more specifically, speech) are reviewed.

2. The musilanguage model

Guided by the belief that music and language possess strong biological similarities as well as important divergences, Steven Brown (2001) developed the Musilanguage model (ML from now on) as an attempt to go beyond the limits of metaphor in the comparison between the two fields. Their most important points of convergence, according to him, are combinatorial syntax and intonational phrasing (Brown, p. 273).

Combinatorial syntax is the capacity of manipulation of discrete units to generate higher-order units that are in some way "meaningful". These units, in both speech and music, are made of sonic events[1]. This ability to create permutations out of a basic set of smaller units is not exclusive of humans, though. Marler (2001) points out that many birds do have this kind of learned ability to reuse and recombine sonic elements in different sequences, as much as we unconsciously combine phonemes to speak words. This "phonocoding" or "phonological syntax" capability (Marler, 2001, p. 36) in birds

[1] However, the identification of such units may be controversial in the musical domain. Linguists generally agree upon the concept of the phoneme as the most basic sound unit responsible for language's basic structure. In an analogous fashion, a considerable large number of researchers take for granted that the "note" would be musical unit correlate of the phoneme. Even if this can be considered true for some specific traditional or popular western music styles, a more accurate look at the diversity of musical "facts" around the world can put this assumption under serious questionings (e.g., one can think of the minor importance of the traditional "note" concept in such different musics such as Gyorgy Ligeti's work, on the one hand, and the Asian Tuva throat-singers, on the other hand). In the same way, the definition of other higher-level units such as the musical "phrase" — which Brown describes as "one of the basic units of structure and function [for both music and language]" (p. 273) — may not be so self-evident and therefore would benefit from a broader understanding of diverse musical manifestations that are not primarily note-structured.

does not mean, however, that they are actually "speaking": "Song sequences are not meaningfully distinct, in the referential sense; they are rich in affective content, but lacking in symbolic content" (Marler, 2001, p. 40)[2]. Interestingly enough, phonocoding is present in birds but not in non-human primates, what leads that author to consider bird sonic behavior closer to human music than that of our direct relatives apes and monkeys[3].

Besides of the ability of using sound elements as building blocks for complex combinatorial structures, Brown refers to intonational phrasing as the second major point of convergence between language and music. This is defined as the "modulation of the basic acoustic properties of combinatorially organized phrases for the purposes of conveying emphasis, emotional state, and emotive meaning" (Brown, 2001, p. 273)[4]. It can take place at a local and at a global level, a differentiation mostly based on the time span of modulations. Small modulations in short intervals of time are seen as local and affecting individual elements within the phrase structure, whereas broader modulations affecting larger time intervals are analyzed as global expressive phrasing.

These shared aspects of music and language (combinatorial syntax and expressive or intonational phrasing) develop into two higher levels, namely, the phonological and the meaning level. The first is basically the phonocoding capability explained above. In this phonological level, Brown (p. 275) draws a parallel between language and music by

---

[2] There is an ongoing debate in the literature about the existence of some degree of referentiality in food and alarm calls of some birds and mammals. The vervet monkeys from Africa, for example, use specific calls to refer to specific dangers such as snakes, leopards and eagles (Marler, 2001, pp. 32-36). To date, however, the majority of animal calls is not seen as symbolically meaningful, in spite of the significant findings above mentioned.

[3] Phonocoding here refers solely to the capacity of making permutations using discrete sound units, which is commonly found in birds but not in primates. The problem of referentiality of sounds does not depend directly on this combinatorial ability.

[4] However, the concept of "emotive meaning" is never clearly defined in Brown's text.

relating "phonemes" to "pitches" and "morphemes" to "motifs".[5] On the other hand, the meaning level is, broadly speaking, the "symbolic content" that Marler referred to, but Brown elaborates this concept in a more detailed manner. The idea of sound meaning as a continuum oscillating between two poles (referential versus emotive meaning) is central to the ML model. Music and language would differ more in emphasis than in kind (Brown, p. 275), and mixed manifestations like poetry or verbal songs would appear in the middle of this continuum. Referentiality would be characteristic of language, while emotivity would be characteristic of music. Again, it is not clear for us the precise definition of the "emotive meaning" concept. One could draw a parallel of this with Umberto Eco's semiotic approach in his book The Open Work (1989), in which he posits that daily language tends to function best by avoiding ambiguity, while in arts ambiguity is desired and even expected (poetry being an interesting case where language itself has to be "forced" towards ambiguity). Thus here we are assuming that emotive meaning — as one end of the continuum proposed by Brown — is related to the notion of ambiguity and non-referentiality (or non-denotative meaning). Finally, the ML model is based on the assumption that "common features of these two systems are neither musical nor linguistic but *musilinguistic*, and that these properties evolved first" (Brown, p. 277). This musilinguistic stage is divided into two phases: in the first one, lexical tone plays a fundamental role, being the initial step for the discreteness of pitch units within broad semantic meaningful vocalizations. In a second step, combinatorial phrase formation and intonational (expressive) phrasing are developed. After this stage, music and language

[5] In spite of our disagreement as to this "note-based" hierarchy, let's let it aside for now to follow the author's ideas as they are. See footnote 1.

would eventually split into specialized human activities, without losing these basic shared ancestor features while giving rise to other exclusively musical or linguistic features.

2. <u>Evolutionary musicology</u>

The ML model is linked to a body of research identified as the evolutionary theories of music. There are many other evolutionary approaches trying to understand the origins of music. Huron (2001) identifies up to eight broad lines of such theories, ranging from sexual selection to motor skill development and social cohesion, among others. Some of these approaches are not necessarily mutually exclusive, but rather they differ in focusing more on one or another possible explanation as to why and how music would have evolved. Other approaches are more controversial. For example, Steven Pinker sustains that music has actually no intrinsic evolutionary value, being merely some sort of "by-product" of other important adaptive traits (Cross, 2001). Another example in a different direction is Geoffrey Miller's theory of evolution of music through sexual selection (Miller, 2000; 2001a; 2001b): he departs from Darwin to develop the hypothesis that, ultimately, "music evolved and continues to function as a courtship display, mostly broadcast by young males to attract females" (Miller, 2001b, p. 354).[6]

However, one central issue for further development of theories of this kind is what types of evidence one can present to support them. Huron (2001) lists four general types,

---

[6] Although the two authors (Pinker and Miller) present many valid arguments in defense of their theories, we agree with Ian Cross when he states that "[such] theories suffer from an attribute that disqualifies their conclusions from serious consideration, that of ethnocentricity. (…) [They focus] on only one dimension of music (…): that of music's inefficacy in any domain other than the individually hedonistic" (Cross, 2001). This "hedonic ends" view would be a reflection of the "subsumption of music into the capitalist economy as a tradable and consumable commodity", after the development of "recording technology together with the reification of intellectual property and the globalization" (Cross, 2001). Underpinning this criticism is the study of the diversity of music around the world, the variability of its uses and the different social values it can have within and beyond western culture. The question of how to reconcile nature and culture is still the center of the debate: what in music can be explained in general and scientific terms and what should be studied under the lights of particular cultures?

namely: Genetic, Neurological, Ethological and Archaeological evidences. The first one is perhaps the more controversial, since to date there is no proof for the link between music and genes (Huron, 2001, p. 48).

The other kinds of evidence furnish data on a number of different aspects of musical activity:

- Archaeological findings that prove the existence of music dating back to the most ancients human settlements (Huron, 2001, p. 48);
- The ubiquity of music across all human cultures may be an anthropological evidence that supports an evolutionary argument (Huron, 2001, p. 50);
- Empirical data describing brain structures responsible for music cognition can provide important insights about how music works in humans' minds, and what is innate and what is the result of learning. Experimental comparisons can help in further elucidation as of what is common and what is specific in music and language processing.

One interesting theoretical framework for comparative studies with animals has been proposed by Hauser et al. (2003). One of the major advantages of comparative studies is that, with animals, one can control in detail the degree of lifetime musical exposure to an extent not possible with humans. Also, Hauser argues that if any perceptual feature supposedly specific to music is found in nonhuman animals, it is unlikely that it was part of an adaptation for music, "and must instead represent a capacity that evolved for more general auditory scene analysis" (Hauser, 2003, p. 665).[7] For example, an experiment is

---

[7] Such arguments can relativize the view of music as a mere evolutionary "by-product" of other cognitive capacities. Music might have a real evolutionary value while at the some time it could have taken advantage of other more general, already established perceptual features: "Human and nonhuman animals thus encode emotional information in their vocalizations and have perceptual systems that are designed to

described in which rhesus monkeys showed octave generalization (Hauser, 2003, pp. 665-666).[8] This would mean that octave generalization is an innate perceptual constraint not exclusive to humans. This kind of experiment represents a valuable path for future research not only to achieve a better understanding of the specificity of musical activities but also in the disentangling of the "nature versus culture" problem.

Another evolutionary approach is the social bonding hypothesis. According to this view, one of the main adaptive values of music would be its capacity of creating or strengthening social cohesion among groups. Huron (2001) enumerates several arguments and evidence in support of this view. One example is his comparison between two different mental disorders: Williams Syndrome and Asperger Autism. He points out the link of symptoms such as low sociability and low musicality in the latter, as opposed to high sociability and musicality in the former (Huron, 2001, pp. 54-55). Another example is the idea of music as a large-scale bonding mechanism more efficacious (for groups) than the rather interpersonal "grooming, gossiping, courting and conspiring". Music, more than language, would have "allowed humans to live in larger groups with their attendant complex social relations" (Huron, 2001, p. 53).

To sum up, so far we have discussed a) the musilanguage model, which tries to explain <u>how</u> music and language would be related in terms of a common origin, and b) other evolutionary accounts that try to understand the <u>why</u> music would have evolved and what could be its adaptive value. After reviewing some of these different accounts, we believe that the most promising theories are those pointing to an <u>integrative</u> view of the

---

respond appropriately to such signals. Given its evolutionary ancestry, our music faculty may well have co-opted this mechanism for use in music, even if it did not evolve for this function" (Hauser, 2003, p. 666).

[8] However, it is not clear whether the experiment controlled or not the monkeys' previous exposure to western music. This could have affected the results, supposing that octave generalization can be learned someway (Hauser, 2003, p. 666).

multiple functions that music may have. In other words, it seems to us that future research should try to integrate the multiplicity of functional levels often displayed by musical activities in an unifying theoretical framework, leading to what Brown has called "multilevel selection models" (Brown, 2001, p. 297).

3. Cognitive Experimental Approach

In their text "Is music autonomous from language? A neuropsychological approach", Patel and Peretz (2001) conclude, "The evidence reviewed (…) suggests that music and language are not independent mental faculties, but labels for complex sets of processes, some of which are shared and some different. Neuropsychology allows the empirical delineation of the boundaries between these domains, as well as an exploration of their overlap" (p. 208). It is this kind of research that can be helpful for further elaboration of theories that account for common origins of language and music, such as the musilanguage model described earlier. Let us now examine some specific experiments that have been conduced in this direction.

3.1. *Rhythm in language and music*. Patel and Daniele (2003) have designed an experiment in which they apply a quantitative measure to compare musical and speech rhythmic patterns. They have compared English and French languages with English and French music themes from a specific era of music history. The hypothesis was that global patterns of rhythmic organization in music could be related to the composer's native language rhythmic structure.

French and English are respectively recognized as a "syllable-timed" language and a "stress-timed language". This basically means that there is quantitative rhythmic difference between these types of language. These differences would exist partly because of variability of vowel durations, and indeed there is recent evidence for this claim: a special kind of measurement has been used by linguists to compare rhythm in English and French utterances, and the data show that significant differences were found (see Patel and Daniele, 2003, for details).

Based on that, Patel and Daniele (2003) designed their experiment using English and French instrumental music. They applied the same quantitative measure of speech rhythm taken from linguistics experiments[9] to the music of 16 composers from those two countries. Their results showed that "English and French classical music have significantly different nPVI values, and that this difference is in the same direction as that observed for language" (p. B42). Thus the claim is that the particular rhythmic organization of a composer's native language can influence him in the rhythmic structure of his compositions.

While it is appealing, however, this idea should be seen with caution. The assumption that notes correspond to vowels was the basis of their experiment ("Vowels form the core of syllables, which can in turn be compared to musical tones", p. B37). This association is doubtful. The use of "note-units" in music is far more elastic and flexible than the use of vowels in speech. Also, not always the single note can be taken as the basic unit in determining the rhythmic profile of a musical passage. Besides that, speech is constituted of sequences of vowels and consonants, and the organization and interaction of these different sound qualities is an integral part of its rhythmic structure. By assuming that all notes in a string of tones are like "vowels", there is necessarily some important

[9] The measurement is called "normalized Pairwise Variability Index" (nPVI). Refer to Patel and Daniele, 2001, p. B37 for more detailed information.

component being missed. This could lead to mistaken comparisons between supposedly similar experiments.

Anyway, the experiment is an important attempt in the direction of concrete comparisons between aspects of music and language, and it raises relevant questions and directions for future research.

3.2. *Words, music and brain imaging*. Besson and Schön (2001) conduced experiments in order to detect event-related brain potentials (ERPs) elicited by different degrees of musical and linguistic events semantically unexpected or incongruent within a given context. It was known from linguistic experiments that unexpected words ending a sentence elicit in the brain a negative component called N400. By testing musicians and non-musicians with familiar and unfamiliar melodies ending with or without "unexpected notes", they found that musical incongruity within a tonal context causes a P600 component. This is seen as a very similar behavior in relation to the linguistic N400 component[10]. Further experiments combining melody and words (by using opera themes) showed that incongruous words, even when they are being sung, do elicit a N400 component, while incongruous notes (no matter the "congruity" of the word accompanying it) always elicited a P600 component.[11]

According to the authors, "this finding provides 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" (Besson and Schön,

[10] P and N mean, respectively, "positive" and "negative", referring to voltage variations detected in the brain. The numbers 600 and 400 refers to how many milliseconds it took to the variation peak after the target onset (the incongruous note or word). The meaning of this positive-negative qualitative difference still remains unclear (Besson and Schön, 2001).

[11] All possible permutations were equally tested during the experiment: (1) congruous final word sung "in tune"; (2) congruous final word "out of tune"; (3) incongruous final word sung "in tune" and (4) incongruous final word sung "out of tune". See Bessson and Schön (2001), p. 246, for details.

2001, p. 246). It may be possible that other levels of language and music processing also share general cognitive processes. This would be the case for detecting syntactic and harmonic violations.[12] The results of studies comparing ERPs associated to grammatically unexpected violations within sequences (language) and presentation of chords from distant tonalities within a simple sequence (music) showed that "the effects of violation of syntactic and harmonic expectancies were not significantly different" (Besson and Schön, 2001, p. 250).[13]

Another interesting experiment was made by Besson and Schön to study the effect of temporal ruptures in musical and spoken sequences. Familiar and unfamiliar phrases were presented. Some of them had the final note or word delayed by 600 ms, thus breaking the expected flow of events in listeners. Different and complementary brain imaging methods were used to get a good combination of temporal resolution and good localization of brain activations. Briefly, results have shown that "similar brain areas were activated by temporal violations in both language and music", leading the authors to conclude that "taken together our results suggest that processing temporal information in both language and music relies on general cognitive mechanisms" (Besson and Schön, 2001, p. 254).

3.3. *Designing an experiment.* Following up Hauser's (2003) speculations on the human-nonhuman comparative approach, we have designed the following experiment.

Background:

The experiment conducted by Ramus and colleagues (2000) suggests that there are interesting similarities in the perceptual mechanism of newborns humans and tamarins:

[12] Tonal harmony is taken as "musical syntax" here. One can suppose that similar results would be achieved for any other kind of non-tonal musical structure ("syntax"), provided that the listeners have learned or internalized it someway. This hypothesis, however, remains to be verified.

[13] This experiment was made by Aniruddh Patel and collaborators, and summarized in Besson and Schön's article (2001, p. 250).

under experimental conditions, both species were able to discriminate two rhythmically distinct languages (Dutch and Japanese). This implies that some aspect of speech processing might not be unique to humans. Speech rhythm is likely to be the key for explaining such results, and "although human infants may be equipped with a capacity to discriminate languages on the basis of rhythmic cues, the presence of this capacity in nonhuman primates that lack language suggests that it evolved for more general auditory purposes" (Hauser, 2003, p. 667).

Design:

Our experiment is based on Ramus' experiment (RE), and it points to three different directions.

1) Replicate RE eliminating the primates' possible previous exposure to human language. We propose to test newborn humans and young tamarins. The nonhuman primates in RE were not young and were not controlled in terms of previous exposure to human language. This could lead to some distortion in the comparison of the "innateness" of the studied traits. The newborn or young tamarins should have no exposure to human language, nor should their mothers have such exposure during pregnancy. Let us refer to this group of human infants and newborn tamarins as 1A and 1B.

2) With another group of subjects (2A and 2B, same characteristics of group 1), repeat RE but now using filtered (narrow band) speech samples as stimulus. Shannon (1995) showed that good levels of speech recognition could be maintained with as little as four bands of modulated noise. We propose to use this method to process the speech samples from Dutch and Japanese languages, and use the result as the stimuli for groups 2A and 2B. By narrowing down the amount of information in speech to its basic rhythmic

component with little spectral cues, we expect to determine whether or not rhythm is the determinant factor responsible for language discrimination in the two species, without significant help of timbre and pitch elements.

3) Finally, with a similar group of human and nonhuman subjects (3A and 3B), we would use other kinds of rhythmic patterns not derived from or directly related to speech. In order to keep the concept of discriminating two significantly different rhythmic organizations in an acoustic signal, we propose to compose short samples of musical sequences based on the *regularity — irregularity principle*. The concepts of periodicity and aperiodicity can be used in order to differentiate two broad classes of rhythmic organization[14]. Short rhythmic sequences can be composed in such a way that they stress or not a periodic underlying pulse. An "aperiodic sequence", however, does not mean a non-structured sequence — there are many ways of organizing a musical discourse non-periodically.[15] Therefore, we would have two classes of organized musical patterns representing both ends of the continuum "regular-irregular". A single percussive, non-pitched timbre should be used to record the rhythmic sequences to be used as stimuli.

General Methodology:

* Habituation procedure: human infants and tamarins are familiarized with one class of stimuli (one language or one kind of rhythmic pattern). After the habituation period (when subjects lose interest in the stimuli), the stimulus type is changed. Their responses

[14] And also spectral organization, as it is the case in electro-acoustic music composition, where the composer is able to work in the micro-world of timbre structure.

[15] Refer to Messiaen, O. (1956). *The technique of my musical language*. (Translated by John Satterfield). Paris, A. Leduc, and also Gramani, J. (1992), *Rítmica*. São Paulo, Ed.Perspectiva. Both authors show a number of different ways of exploring aperiodicity musically.

are measured (sucking rate in newborn humans and tamarins[16]) and compared within group and between groups. Higher sucking rates indicate a renewed interest for the acoustic signal (novelty response).

Questions & hypotheses:

We know from RE that both species can discriminate sentences from two rhythmically different languages. The first question is: to what extent are rhythmic cues really crucial to this discrimination? The second question is: supposing both species share a common innate mechanism for dealing with rhythmic perception, can they discriminate different types of musically organized rhythmic structures (not directly related to speech)?

We present below some possible results that can come out of the experiment:

* If the "RE replication" group (1A and 1B) does make positive recognitions of the different languages: this would confirm that this capacity of language identification in both humans and tamarins is likely to be innate. If the young tamarins in this group were unable to make positive recognitions, this would reveal a possible flaw in RE, indicating that previous exposure to human language had actually influenced the older primates' perception.

* If the "filtered speech" group (2A and 2B) does make positive recognitions of different signals: this would suggest that there is an innate capacity of identification of speech rhythmic patterns in both humans and tamarins. If this group shows no discrimination, one could assume that rhythmic cues are not enough to account for the language identification capacity found in the first group.

---

[16] Can the baby tamarins' reactions be measured in terms of sucking rate? In RE, since they were older tamarins, the head-turning method was used. We lack the technical knowledge of working with animals to assess the viability of this hypothesis.

* If the “musical” group (3A and 3B) does make positive recognition of different stimuli: this would suggest that the attention to rhythmic cues is NOT language-specific, but rather it works on basic aspects of auditory temporal perception (thus encompassing music and language). If this group shows no discrimination, this would mean that either a) the capacity under study is hard wired for speech-like signals (problem: why would primates be able to perceive speech if they don’t have language? Then one would have to look for similarities between speech and primate vocalizations) or b) our test was flawed by some bad experimental design.

5. References

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

Besson, M., & Schön, D. (2001). Comparison between language and music. In Zatorre R., & Peretz I. (Eds.), *The Biological Foundations of Music* (pp. 232-256). New York: The New York Academy of Sciences.

Brown, Steven (2001). The “musilanguage” model of music evolution. In Nils L. Wallin et. alt. (Eds.), *The Origins of Music* (pp. 271-301). Cambridge: MIT Press.

Cross, Ian (2001). Music, cognition, culture and evolution. *Annals of the New York Academy of Sciences*, Vol. 930, pp. 28-42.

Cross, Ian (2003). Music and biocultural evolution. In Clayton, M. et al. (Eds.), *The Cultural Study of Music: a Critical Introduction* (pp. 19-30). New York/London: Routledge.

Drake, C., & Bertrand, D. (2001). The quest for universals in temporal processing in music. In Zatorre R., & Peretz I. (Eds.), *The Biological Foundations of Music* (pp. 17-27). New York: The New York Academy of Sciences.

Eco, U. (1989). The Open Work (A. Cancogni, Trans.). Cambridge, MA: Harvard University Press.

Hauser, M. & McDermott, J. (2003), The evolution of music faculty: a comparative perspective. *Nature Neuroscience*, Vol. 6(7), pp. 663-668.

Huron, David (2001). Is music an evolutionary adaptation? In Zatorre R., & Peretz I. (Eds.), *The Biological Foundations of Music* (pp. 43-61). New York: The New York Academy of Sciences.

Marler, Peter (2001). Origins of music and speech: insights from animals. In Nils L. Wallin et. alt. (Eds.), *The Origins of Music* (pp. 31-48). Cambridge: MIT Press.

Miller, Geoffrey (2000). Mental traits as fitness indicators: expanding evolutionary psychology's adaptationism. At http://www.unm.edu/~psych/faculty/mental_traits.htm

Miller, Geoffrey (2001a). Aesthetic fitness: how sexual selection shaped artistic virtuosity as a fitness indicator and aesthetic preferences as mate choice criteria. *Bulletin of Psychology and the Arts*, 2(1), pp. 20-25.

Miller, Geoffrey (2001b). Evolution of human music through sexual selection. In Nils L. Wallin et alt. (Eds.), *The Origins of Music* (pp. 329-360). Cambridge: MIT Press.

Molino, Jean (2001). Toward an evolutionary theory of music and language . In Nils L. Wallin et alt. (Eds.), *The Origins of Music* (pp. 165-176). Cambridge: MIT Press.

Patel, A. D. & Peretz, I. (1997). Is music autonomous from language? A neuropsychological appraisal. In Deliège, I. (Ed.), *Perception and Cognition of Music* (pp. 191-215). Hove, East Sussex Psychology Press.

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

Patel, A. D. & Daniele, J. R. (2003). An empirical comparison of rhythm in language and music. *Cognition*, 87, B35-B45.

Ramus, F. et al. (2000). Language discrimination by human newborns and cotton-top tamarins (pp. 349-351). *Science*, 288.

Shannon, R. V. et al. (1995). Speech recognition with primarily temporal cues (pp. 303-304). *Science* 270.

Ujhelyi, Maria (2001). Social organization as a factor in the origins of language and music. In Nils L. Wallin et alt. (Eds.), *The Origins of Music* (pp. 125-134). Cambridge: MIT Press.