MUSIC AND LANGUAGE: CURRENT STATE OF RESEARCH

Shuo Zhang
Department of Music
University of Pittsburgh

ABSTRACT: This paper surveys literature on the studies that relate music to language, cognitively and structurally. Highly interdisciplinary in nature, the research in each area seeks to approach the question of music and language from different angles. In this paper I investigate the current state of research by answering these primary questions: 1) What are the types of approaches taken in overall research designs? 2) What has been done in each of these major domains of comparisons, as identified by Patel (2008): sound elements (pitch and timbre), rhythm, melody, syntax, and meaning? 3) What are the answers that we have now to the central question?

KEYWORDS: music; language; cognition; linguistics; empirical

0 Introduction

Recent years have witnessed a growing interest in the comparative research on music and language, ranging from structural/formal approaches (e.g. Lerdahl and Jackendoff ,1983; Pesetsky et al., 2009) to empirical approaches (e.g. Besson and Friedrici,1998).[1] The breadth of the types of research that can possibly contribute to the understanding of this subject is vast, as exemplified by Patel's book *Music, Language, and the Brain* (2008), an attempt to draw comparisons between music and language by synthesizing a huge body of literature including those from musicology, ethnomusicology, music theory, linguistics, cognitive science, neuroscience, and evolutionary biology.

Methodologically, two approaches are commonly taken. Linguistics theory takes an approach of formal analysis, explicitly articulated through abstract rules derived from analysis. Empirical methods are employed in a number of disciplines (e.g., cognitive psychology, psycholinguistics, empirical musicology, music therapy), which seek to approach the question of music and language cognition from different angles. No matter what approach is taken, the central question remains, "what perceptual and cognitive

[1] Other approaches, such as the anthropological approach to music and language (exemplified by Feld and Fox 1994), are not discussed in this paper in general.

operations are specific to language, and which are also involved when processing other types of well-constructed and organized signals, such as music?" (Besson&Friederici 1998)

Scope

Two kinds of research in music-language comparison are identified for the discussion of the current paper: (1) Directly comparing aspect(s) of music and language structure or cognition in one single study under one unified theoretical and analytical framework; and (2) Taking what is known from existing research on linguistic theory and music theory (or independent empirical research in each domain) and make comparisons on variously levels. It is worth noticing that in Patel's book (2008), the majority of discussion belongs to type (2) in nature. An examination of Patel's bibliography also reveals that the majority of literature cited belongs to independent research in either domains of linguistics (including psycholinguistics) or music (including music cognition), whereas the amount of type (1) research only occupies a small percentage. This is understandable due to the fairly recent interest in comparing the structure and cognition of music and language, and much more research are expected to be done (which is why Patel's book also contains more than 30 proposals on how such research could be conceived) (Schon, 2009:287). Thus it is the intention of this paper to focus on the current state of type (1) research.

Three questions of particular interest concerning type (1) research are: [i] What are the types of approach taken, and what are the levels of comparisons? [ii] What has been done in each of these major domains, as identified by Patel (2008): sound elements (pitch and timbre), rhythm, melody, syntax, and meaning? [iii] What are the answers that we have now to the central question?

This paper reflects the status of the research by taking the following core journals of music cognition as main sources, where the comparison of music and language occurs most intensely: *Music Perception; Musicae Scientiae; Psychology of Music; Psychomusicology*; and *Empirical Musicology Review*.

I Music and Language: Parameters

The study of music and language, much like other domains of human cognition, has been benefited from a wide variety of approaches, each taking a different angle to reveal part of the puzzle. It is suffice to say that such multi-disciplinary approach is necessary in that each angle can compensate, solidify, and test the conclusions derived from others, due to the often-different innate assumptions and logics of the respective research designs and methodologies. For instance, Patel (2008: 72) stressed the possibly similar mechanism in sound category learning in music and language despite the fact that there are good reasons to believe that the brain treats spoken and musical sound systems differently, according to evidences such as the dissociation for perceiving spoken versus musical sounds after brain damage. In other words, "even though the sound category representations in the two domains, once learned, do not completely overlap in terms of their location in the brain", "whether or not similar mechanisms are used to *create* these representations is an orthogonal question."(ibid:73)

On the other hand, it is often necessary to specify exactly on which level the comparison is being made. One attribute that music and language shares is that they both involve multiple levels of representation: phonological prosodic, lexical, semantic, syntactic, and pragmatic in the case of language and rhythmic, melodic, and harmonic in the case of music (Besson & Friederici 1998:2). The mechanisms involved on each level are often different, and the understanding of each level is crucial to the conception of the whole picture. In a study on musical rhythm and linguistic rhythm and their role in human evolution (Patel 2006), for example, Patel comments on the articles by Justus and Hutsler (2005) and McDermott and Hauser (2005), who suggest that musical pitch perception can be explained without invoking natural selection for music. Patel believes that these authors leave the issue of musical rhythm largely unexplored and proposed how various issues might be addressed by extending their conceptual approach to the domain of rhythm. In addition, contrary to the previously firmly held belief of hemispheric asymmetries for music and language, there is increasing evidence suggesting the bilateral representation in the auditory cortex in the brain for both music and language sound categories (Patel 2008:73), making the various levels of comparison more significant (i.e., which level is represented at which hemisphere, or involved in both hemispheres).

In light of these two perspectives, I summarize the parameters involved in current research literature. These parameters, as I define here, include both the approaches taken to compare music and language, and the levels of comparisons, as addressed in the previous paragraphs. By employing and manipulating these parameters and their different roles and functions, the most meaningful comparisons can be made. Parameter 1 through 4 concerns methodology and framework of the research design; parameters 5 and 6 are targeted areas/levels of comparison.

Following the outline of the parameters, the next section will focus on the research literature in each of the music/linguistic parameters of comparison (extracted from Parameters 5 and 6 combined): Tonal functions, temporal functions, syntax, and meaning.[2] Additionally I also discuss the area of research that has received much attention recently, links between the general musical training and specific linguistic abilities.

Parameter 1 Subject Tested:
(1) Children/adult with musical or linguistic deficiencies/cognitive disability/brain injuries, such as text dyslexsia and aphasia;
(2) Normal children/adult with no identified cognitive deficiencies. Often further divided into two subgroups:
  a. Musically trained;
  b. Musically untrained;[3]

Parameter 2 Time Frame
(1) Diachronic: Developmental/Acquisition/Process of Learning;
(2) Synchronic: Subject carrying out perception tasks with skills already acquired;

Parameter 3 Cognitive Domains

[2] The usage of tonal and temporal functions here follows the classification of musical parameters by musicologist Charles Seeger (1977). While Seeger's writing did not explicitly concern music and language perception, this conceptualization works effectively for the current purposes.
[3] In some cases the distinction between musician and non-musician is used.

(1) Aoustic: Production and perception of acoustic signals;
(2) Aoustic-Visual Interface: such as processes involving music reading and language reading;

Parameter 4 Disciplinary-Specific Methodology/Tools (such as empirical method; brain imaging with fMRI and other technological tools; MIT in music therapy; formal analysis in linguistics; etc.)

Parameter 5 Linguistic Parameters
(1) Phonetics/Phonology (including sound categories and prosodic structures);
(2) Syntax;
(3) Semantics;

Parameter 6 Musical Parameters
(1) Pitch;
(2) Rhythm;
(3) Melody;
(4) Syntax;
(5) Meaning;

II Music and Language: Research in Five Areas of Comparison

1. Tonal Functions

Charles Seeger proposed that the tonal functions of music include three components: pitch, loudness, and timbre, of which the first two are simple functions and the latter, a complex function.[4]

Patel (2008) identified two levels of comparison in tonal functions for music and language: sound elements and melody (roughly corresponding to the idea of phonetics and phonology in linguistics). The necessity of this distinction is only apparent after careful consideration of the research literatures. First of all, sound elements include the basic units and smallest building blocks of music and linguistic sound, subdivided into

[4] Simple functions differ from the complex in that they can be represented in a unidimensional fashion, (e.g., the pitch and loudness that can vary by (+, -)[higher, lower] or [louder, softer]), whereas the complex function cannot be measured in a similar manner.

pitch and timbre.[5] On a different hierarchical level, melody can be considered as the meaningful organization of pitch and timbre elements, although it is also much more than the sum of pitch and timbre in that rhythmic and syntactic cues also play crucial roles in the case of the perception of a musical melody, and prosodic structure and pragmatics in the case of linguistic melody. The following section addresses the research on components of tonal functions on these two levels.

(1) Sound Elements

[a] Pitch: Both music and language make use of pitch contrast. Pitch is the most important feature of music. In fact, most cultures in the world make use of rich pitch contrast as their main device in creating music. The pitch also plays a significant role in tone languages, whereas pitch contrast does not infer lexical meanings in non-tonal languages (where intonation plays a role).

Pitch inventories in musical cultures around the world have been investigated in relation to tuning systems and scales since the early 20th century. While little has been theorized about why and how different cultures came to make use of different tuning and scales system, a branch of recent research literature in this area has largely linked pitch perception in music to that in speech sound.

One group of such research concerns the musical illusion known as the tritone paradox (Deutsch, 1991). The tritone paradox occurs when two tones that are related by a half-octave (or tritone) are presented in succession and the tones are constructed in such a way that their pitch classes are clearly defined but their octave placement is ambiguous. Studies have shown that there are large individual differences in how such tone pairs are perceived (e.g. ascending vs. descending), and these differences correlate with the listener's mental representation of pitch range in speech, formed in critical period through

[5] According to Patel (2008), a basic distinction between music and language systems in sound elements is that the former primarily makes use of pitch contrast, whereas the latter primarily makes use of timbre contrast. While this is a valid and clever way to put it, it has received minor criticisms in that the "timbre" used to indicate the contrast among different speech phonemes (such as [i] and [e]) can be confused with the different "timbre" of different speaking voice, i.e., the voice quality of different individuals' speech sound (Brown 2009). Nevertheless it is an effective way to distinguish the two systems, and also useful in asking questions such as why most music do not choose to make use of timbre as the main contrast device instead of pitch (which Patel also gives an answer in his book) while it is theoretically perfectly possible.

language acquisition. Deutsch and collegues (2004) showed that the perception of tritone paradox could be heavily influenced by speech pattern heard early in life, even for listeners who do not speak their first language fluently.

The same group of researchers also linked data on the perception of absolute pitch in music to one's native language, specifically, the difference in the occurance of AP between tonal and non-tonal language speakers. Absolute pitch (AP) is the ability to identify or produce isolated musical tones (e.g., middle C, concert A) accurately, rapidly, and effortlessly (Deutsch et al., 2004a). Previous studies have linked AP largely to genetics, due to the observed fact that the occurrence of AP is much higher among professional musicians in East Asia than in Western society. Deutsch and colleagues (2004a; 2009), conducting research on another line of evidence, found correlations between the high incidences of AP and the native speakers of tone languages (such as Mandarin Chinese and Vietnamese) instead of genetic heritage (e.g., AP is not higher among the Asian American musicians of Chinese heritage who do not speak a tone language as their native language). Their findings also linked musical AP to linguistic AP, showing evidence that tone language native speakers tend to reproduce speech patterns more precisely pitch-wise at different occasions than non-tone language speakers (although the difference is relatively small and its significance still under debate). Not surprisingly, this line of argument of linguistic AP in tone language speakers carrying over to music received much criticism and the debate is clearly still not over (Schellenberg & Trehub 2008).

Both of the research discussed above, whether proved to be valid or not, reflect the underlying hypothesis that there exists an overt association in mental representation between musical and linguistic pitch perception, often acquired through musical training and language acquisition during the critical period. The investigation of the mutual influence of such abstract representations between music and language perception are thus of general interest in this area. For example, Stegemoller and colleagues (2008) investigated the effect of musical training on vocal production of speech and song, finding evidence that higher levels of musical experience were associated with decreased energy at frequency ratios not corresponding to the 12-tone scale in both speech and song. Taking a developmental approach, Bolduc (2005) showed that musical pitch processing is significantly linked to phonological awareness performance among kindergarten students.

[b] Timbre perception: Patel (2008) points out that pitch contrast is the main device in most musical systems, while timbre contrast is the prominent device in language. This is evidenced by the existence of high degree intelligibility in speech heard in monotones (even in the case of tone languages to some degree) with only articulations preserved (timbre information). Due to this divergence, there are relatively few comparative studies on musical and linguistic timbre perception.

An interesting link exists though, between the timbre of vocables (non-sense made-up words used in a variety of singing styles found in world music)/verbalizations (non-sense words used to represent the different timbres of different musical sounds), and the timbres of actual musical sound found in that culture. Patel and Iverson (2003) conducted an empirical study on the sound symbolism, acoustic and perceptual comparison of speech and drum sounds in North Indian tabla tradition. Analysis revealed that acoustic properties of drum sounds were reflected by a variety of phonetic components of vocables (including spectral centroid, rate of amplitude envelope decay, duration between the releases of consonants in a cluster, fundamental frequency, and the influence of aspiration on the balance of low vs. high frequency energy in a vowel), and the superior performance of 7 non-Hindi speakers unfamiliar with tabla drumming on matching the vocables with their corresponding drum sounds showed that the resemblance between speech and musical sounds in this case is based on shared perception rather than convention.

(2) Melody

I have discussed the necessity of this level of comparison in previous sections. Patel (2008) points out nine areas where human perceptual system concerts a two-dimensional sequence (pitch vs. timbre) into a rich set of perceived relationships (melody), which is one of the reasons why musical melodies tend to stick in listeners' mind more persistently whereas linguistic melodies leave little impression. The nine areas (which is not an exhaustive list) include: grouping structure; melodic contour; beat and meter; intervallic implications; motivic similarity; tonality relations—pitch hierarchies; tonality relations—event hierarchies; tonality relations—implied harmonies; and meta-relations generated among the previously listed relations.

A point of interest in this area concerns the interplay between musical and linguistic melodies in vocal music. For instance, a number of scholars studied the interaction between musical and linguistic tones in various forms of Chinese operas (Yung 1989; Pian 1972). Dresher (2008) examined the relationship between Gregorian and Hebrew Chant. By analyzing the different structures in the interactions among words (prosodic structures, stress), speech intonations, and melodies, she revealed the nature of chanting in contrast with singing, and the reason why it should be considered a form of heightened speech melody rather than musical melody.

The question of whether amusia patients (musical tone deafness, a cognitive deficits in the perception of pitch contour) show deficits in linguistic intonation perception has been long held as a major evidence as to whether and to what extent music and language share neural mechanisms for processing pitch patterns. In a recent study, Patel and colleagues (2008) showed that about 30% of amusics from independent studies (British and French-Canadian) have difficulty discriminating a statement from a question on the basis of a final pitch fall or rise, suggesting that pitch direction perception deficits in amusia can extend to speech. Meanwhile, counter examples of dissociation between the two domains are also reported (Ayotte et al. 2002), leading to a on-going debate that calls for detailed research into the question (such as the identification of the types of amusia, and types of discrimination tasks used in experiments and their implications).

Finally, Patel (2006) reports a recent development of a quantified model that allows comparison between speech melody (either lexical tone melody or intonation melody) and musical melody. The most salient feature of speech melody that differs from musical melody is that speech melody does not employ fixed, discrete pitches like music; rather, the phonological analysis has relied on the contour of F0 (fundamental frequency) to describe speech melody (an alternate description is offered by the abstract phonological tones such as [H] and [L] tones in autosegmental-metrical theories of intonation). This posits fundamental difficulties in comparing speech melody to musical melody. [6] The new prosogram takes the underlying notion that the raw fundamental frequency (F0) contour, although an accurate physical description of the speech signal, is represented differently in a listener's perception. Rather, empirical data suggest that such perception is subject to several perceptual transformations, most notably, [1] segregation of the F0 contour into syllable-sized units due to the rapid spectral and amplitude fluctuations in speech signal, and [2] temporal integration of F0 within the syllable. Based on these ideas, the prosogram converts a sentence's original F0 contour into a sequence of discrete tonal segments, and allows more meaningful comparisons with musical melodies. These converted segments do not necessarily conform to any musical scale (Patel 2006).

Patel illustrates two potential use of the prosogram. For instance, it would allow a more accurate comparison between the linguistic melody and musical melody in a song (such as the case of the Chinese opera mentioned above), and to see if the similarity and difference in the two melodies corresponds to interesting points in musical structure, or the different distribution of such correspondences among songs in different languages and in different musical genres.

It seems also conceivable to me that the prosogram can be potentially used to shed light on some unresolved questions in musical-linguistic illusions, thus deepens our understanding of the pitch perception in music and language. Although no study to my knowledge has been emerged in this new application. For instance, one topic that can be benefited from this tool is the speech-to-song-illusion discovered by Deutsch and colleagues (2008). [7]

[6] Interestingly, there are such European composers, notably Joshua Steele (1700-1791) and Leos Janacek (1854-1928), who were fascinated by the speech melody of other people's talking, and transcribed the melody onto a Western staff notation. Janacek did this for over 30 years, whose transcription was discussed and analyzed in a recent article (Pearl 2006). One problem with the transcription, apparently, is the discrepancies between the pitches used in staff notation and its physical pitch height in reality. Patel argues that the newly developed prosogram is one way to resolve this problem and see interesting discrepancies between the results produced by the prosogram and those by human transcription based on perception.

[7] In this illusion, a spoken phrase is made to be heard convincingly as sung rather than spoken, and this perceptual transformation occurs without altering the signal in any way, or adding any musical context, but simply by repeating the phrase several times over. The illusion is surprising, as it is generally assumed that whether we perceive a phrase as spoken or as sung depends on the physical characteristics of the sound.

(3) Melody-Timbre Interface

A different type of study concerns the interface of various levels and components of tonal functions. This is not surprising given the already discussed rich relations between musical melody and other parameters in music (and potentially in language). An example of such interface is the study by Collister and Huron (2008) on the comparison of word intelligibility in spoken and sung phrases. It illustrates the interaction between the production/perception of phonetic information (timbre contrast) and the linguistic/music melodies that carry such information.

2.Temporal Functions

Although both components of the simple temporal functions, rhythm and tempo are of interest in music perception, research in music and linguistic temporal perception has almost exclusively focused on the rhythm. Even for that part, it was only until recent years that significant progress was made. In this section I focus on the prominent research in recent years regarding comparison and interaction of human music and linguistic rhythm from two different perspectives.

(1) nPVI and rhythmic typology. Virtually most (if not all) significant studies on music and linguistic rhythm in recent years stem from the same methodological framework, developed by Patel and colleagues (Patel and Daniele, 2003, 2003a; Patel, 2003a; Patel, Iversen, and Rosenberg, 2006). This methodology in turn is borrowed from the recent advancement in linguistics that allows a quantitative study of linguistic rhythm typology among languages of the world.

Linguistic rhythm refers to the way language is organized in time, often studied as part of the language prosodic structure. Earlier linguistic theories stressed the idea of isochrony, which attempted to establish the rhythmic typology on the basis of periodicity (similar to the periodicity in musical rhythm). Three types of linguistic rhythm were identified: stress-timed(e.g., English, Dutch, German), syllable-timed(e.g., French, Italian, Spanish), and mora-timed(e.g., Japanese) (Abercormbie, 1967; Pike, 1945). It has been hypothesized that stress-timed languages have equal duration between stresses, while syllable-timed languages show equal duration between syllable onsets (equal duration between moras in the case of mora-timed languages). However, empirical evidence failed to support such hypothesis that such theory has been largely abandoned in phonological research in the 1980s (Roach 1982).

Interestingly, recent research in phonology supported such classification by demonstrating that there is indeed quantitative difference in rhythms between, for example, stress-timed and syllable-timed languages (Low, Grabe and Nolan, 2000; Ramus, Nspor and Mehler, 1999). The key to understand this new approach is to set aside the concept of isochrony (i.e., periodicity in rhythm) and look for alternate regularities and features in different types of linguistic rhythm. This is also stressed by Patel (2008) in considering the comparison between musical and linguistic rhythm, notably "…the mistaken notion that rhythm *is* periodicity, or that rhythm *is* a regular alternation between strong and weak beats, rather than the broader notion of rhythm as systematic temporal, accentual, and phrasal patterning of sound, *whether or not this patterning is periodic*….Many widespread musical forms lack one and/or the other of these features [periodicity and strong-weak beat alternation] yet are rhythmically organized." (Patel 2008:151)[8]

Based on such broad conception, the recent approach to linguistic rhythm thus came to focus instead on the durational patterning of vocalic and intervocalic intervals in speech (Grabe and Low, 2002; Low, Grabe and Nolan, 2000; Ramus, 2002)[9]. Implications of this approach include (in the case of stress vs. syllable-timed languages): [1] stress-timed languages show a greater degree of vowel reduction than syllable-timed languages (Dauer, 1983), thus having a greater variability in vowel durations; [2] stress-timed languages tend to permit more complex syllable structures comparing to syllable-timed languages, thus having a greater variability in the duration of consonant sequences (Ramus et al., 1999).

---

[8] Italicization by Patel.

[9] "Vocalic intervals are vowels and sequences of consecutive vowels, regardless of whether they belong to the same syllable (or word, for that matter) or not. Similarly, inter-vocalic or consonantal intervals are made up of consonants and sequences of consecutive sonsonants." (Ramus 2002)

Patel and colleagues noticed that the quantitative measurement proposed by Grabe and Low (2002) is appropriate for comparisons of rhythmic features between language and music of a particular given culture (in order to investigate the possible relationship between them). This is the "Normalized Pairwise Variability Index", also known as nPVI, indicating the durational variability within a given language, or, to put it more plainly, "measures the degree of durational contrast between successive elements in a sequence" (Patel, Iversen, and Rosenberg, 2006).[10]

Setting out to test the long held popular notion among certain scholars that speech patterns are reflected in the instrumental music of a given culture (Abraham, 1974; Hall, 1953), Patel and Daniele applied nPVI to the durations of notes in instrumental classical themes from England and France (representing typical stress-timed and syllable-timed languages), and found that English music had a significantly higher nPVI than French music (being consistent with the higher nPVI for British English vs. French)[11]. Subsequent works by others showed that this finding could be generalized to a much broader historical sample of composers from England and France (Huron and Ollen, 2003), though exceptions are observed (Patel 2003a). Hannon (2009) also showed that after a period of brief training, listeners can successfully classify instrumental arrangement of folk songs as from Britain or France only according to the rhythmic typology cues (i.e., different in nPVI, which is in turn consistent with different linguistic rhythms of French and British English). Similar studies using nPVI are also conducted in Korea (Slobodian 2008; Iverson 2008) and Japan (Sadakata et al. 2004), whose results are generally in consistent with the parallel between music and language rhythmic typology, while the parallels are to a less extent than expected (Sadakata et al.2004).

[10] Patel and colleagues also discussed the meaning of nPVI as "durational contrastiveness" despite its name "durational variability" in relation to the computation of nPVI. In a nutshell, the nPVI "is a contrastiveness index and is quite different from measures of overall variability (such as the standard deviation) [within a give language]). "((Patel , Iversen, and Rosenberg, 2006: 3035). Appendix I gives the computation of the nPVI.

[11] It is often necessary to point to the exact variation of a language when talking about rhythmic typology. For instance, while British English is stress-timed, Jamaican and Singapore English are considered syllable-timed. This is due to the role the rhythm plays in the prosodic structure of speech. One example that illustrates this property of rhythm in language is its role in second language acquisition. Without correct linguistic rhythm, a second language learner can still produce intelligible utterances; however, for native listeners, the misplacement in rhythmic structure is a significant factor in creating foreign accents.

In sum, the application of nPVI in linking musical to linguistic rhythmic typology allows quantified comparative studies to be made on a valid theoretical basis that seems to be fruitful so far, especially in the case of French vs. British music and languages. However, much still needs to be done to broaden the scope of the application, and possible difficulties are also present in deciding the historical and cultural influences/interactions in shaping the rhythmic characters in the music of a given culture, not to mention the great diversity of music existing in any one given culture. Before these issues can be carefully and effectively dealt with, the universality of this hypothesis cannot be attested.

(2) Music rhythm, linguistic rhythm, and human evolution. Darwin (1871) was the first scholar to propose the idea that human minds have been shaped by natural selection for music. While recent years have seen the renewed interest in this idea, skepticism also arises as to whether music is a byproduct of other cognitive functions that have more clearly adaptive values such as language. Hypothesis that argues a "musilanguage" stage in human evolution (Brown, 2000; Mithen, 2005; Jordania, 2006) also emerged in recent years.

Following the discussion on the role of pitch perception in music and language in human evolution by Justus and Hutsler (2005), as well as McDermott and Hauser (2005), Patel analyzed the similarity and differences in music and language rhythm that can or cannot be explained as a byproduct of other cognitive domains (such as language). According to Patel (2006a), while perceptual grouping can be viewed as derived from linguistic prosodic grouping abilities, the feature of periodicity (discussed earlier in this paper) seems to be unique to musical rhythm, thus indicating domain specificity. More specifically, Patel stressed the phenomenon of beat perception and synchronization (BPS), the ability to move or tap according to the beat perception of a given tempo, and argued that this aspect of rhythm appears to be unique to music (Patel, 2006a: 100). [12]

[12] Following the discussion of BPS in terms of innateness, domain-specificity, and human-specificity, Patel proposed "vocal learning and rhythmic synchronization hypothesis", which predicts the impossibility to teach non-human primates to synchronize to a musical beat. Further articles were published (Patel et al., 2009) upon the discovery of a cockatoo (a type of bird), which is able to synchronize to the musical beat on various tempos.

3.Syntax

For a long time in the 20^{th} century, the link between musical and linguistic syntax remained in the application of linguistic analytical methodologies (including quasi-linguistic methodologies) to music analysis, most notably, the now-still-widely-taught Schenkerian analysis and semiotic analysis (which actually deals with musical syntax rather than semantics). The most representative theory and analytic framework inspired by generative linguistics is the Generative Theory of Tonal Music (GTTM) proposed by Lerdahl and Jackendoff (1983), which celebrated its 25 years' anniversary by a volume of *Music Perception* devoted to the subject.

It is noticeable that although the GTTM was inspired by generative grammar in linguistics, the authors did not explicitly compare musical syntax with linguistic syntax. In fact, they were largely skeptical of such comparisons (see later discussions in this paper for the recent article of Jackendoff). Indeed, even the use of tree structure shared by both GTTM and linguistic syntactic theory did not resemble each other on various levels (e.g., the absence of the constituency relationships in musical syntactic trees). Patel (2008) points out several differences between music and linguistic syntax representations, such as the presence of grammatical categories in language (i.e., nouns, verbs, etc.), long distance dependencies in languages, and the tolerance of the interpretation of syntactic ambiguities by listeners of music and language. Formal similarities are also identified, including the hierarchical structure in music and language, and the logical structure in the two domains (Patel 2008: 264-267).[13]

Recent debate has focused on the question of whether there is shared or domain-specific processing of music and linguistic syntax from research in neuroscience. Contradictory evidences of the debate include on the one hand, the well-documented disassociations between the two (such as individuals with normal hearing showing impairment of syntactic processing in one domain while remaining intact in the other, e.g., aphasia without amusia, or vice versa), and on the other, neuroimaging data pointing to the overlap in the processing of linguistic and musical syntax.

[13] One study by Jonaitis and Saffran (2009) shows the similarity of music and linguistic syntactic acquisition by studying the role of serial statistics (statistical regularities underlying musical harmony and language learning.)

Drawing research literature from psycholinguistics (Caplan and Waters, 1999; Ullman, 2001), Patel (2003) proposed a possible resolution to this paradox: a conceptual distinction between syntactic representation and syntactic processing in brain structure. A key idea of this "dual-system approach" is that "at least some of the processes involved in syntactic comprehension rely on brain areas separate from those areas where syntactic representation reside." (Patel 2003: 676) In supporting of his hypothesis, Patel also cites linguistic and music theory, namely, Dependency Locality Theory (DLT) (Gibson 1998; 2000) and Tonal Pitch Theory (TPS) (Lerdahl, 2001), both of which posit that integration is influenced by the distance between X and Y [two incoming elements in a evolving musical or linguistic structure] in an abstract cognitive space (Patel 2003: 678). In sum, Patel's proposal resolves the evidences of dissociation with association in the two domains by pointing to the distinct processing/mechanism localities between the storage/representation of the syntactic information and the actual processing of the syntax (which makes use of these stored representation). Thus when the brain is damaged, it is likely that the impairment is located in the storage of one domain, while the processing mechanism remains intact.

4.Meaning

The comparative study on musical and linguistic meaning is relatively rare, given the salient difference between the two systems. Traditionally this area falls into the field of music semioticians (Nattiez 1990) or music philosophers (Seeger 1975). Recently, inspired by the concept of linguistic prosody, Palmer and Hutchins (2006) explored the significance of "music prosody" in relation to musical meaning.

Since music is a non-referential system, comparisons of musical and linguistic meanings in recent empirical research often concern the affective response depending on the acoustic cues employed in music and speech. For instance, Ilie and Thompson (2006) used a three-dimensional model of affect to compare affective consequences of manipulating intensity, rate, and pitch height in music and speech. Their results showed that both faster music and speech were judged as having greater energy than slower music and speech, whereas the pitch height showed opposite affective response for speech and music. Similarly, Huron (2009) explained the affective consequences of major and minor scales by relating the effect of pitch height to the physiological cues observed in speech.[14]

5. Links Between Musical Training and Linguistic Abilities

Besides the areas of comparisons discussed above, recent years have also seen the rise of the interest in the links between general musical training and the linguistic abilities in specific areas.

One branch of such research comes from music therapy. Schulag, Marchina, and Norton (2008) reported a study of why singing may lead to recovery of expressive language function in the treatment of patients with Broca's Aphasia.[15] They showed that the effect of MIT (Melodic Intonation Therapy) surpassed those of control-treated patient, with significant improvement in propositional speech that generalized to unpracticed words and phrases. The researchers also proposed four possible mechanisms by which MIT's facilitating effect may be achieved: [1] reduction of speed in singing than speech; [2] syllable lengthening; [3] syllable "chunking"(prosodic information helps to group syllables into words and words into phrases); [4] hand tapping. Much of the discussion also concerns the controversial bihemispheric representation and processing of singing and speaking. Similar research is also seen in the literature from speech pathology (Jordania 2006), seen in the well-known fact that stutters do not show problems of fluency when singing.

Patel and co-workers (Patel & Iversen, 2007; Patel, Wong, Foxton, Lochy, & Peretz, 2008) discussed the recent findings in linking musical training to phonetic abilities in language learning. Several previous studies have reported associations between musical ability and accuracy at perceiving phonetic or prosodic contrasts in a native or foreign

[14] Personal communication with David Huron at SMPC 2009 Conference, Indianapolis, IN, Aug.2009.
[15] Aphasia is a condition characterized by either partial or total loss of the ability to communicate verbally, thus a linguistic deficit. As previously discussed, aphasia without amusia is often considered as an evidence for the dissociation between music and language processing.

language (Anvari et al., 2002; Thompson et al., 2004; Magne et al., 2006). In exploring the brain mechanism for this link, a recent study by Wong et al. (2007) suggests that musical experience (based on pitch perception) tunes basic auditory sensory processing circuitry in the brain, which has consequences for language processing. The authors thus discussed the possible extension of the experiment from acoustic pitch perception to linguistic pitch learning, as well as to non-pitch-related linguistic phonemic tasks. Similar studies in this area also linked early music training to linguistic syntactic abilities (Marin 2009) and music training to lexical stress processing (Kolinsky et al., 2009).

Finally, research designs have been developed in recent years to deal with other aspects of cognitive domain in music and language, such as eye movements in music and language reading (Madell and Herbert 2008), memory for verbal and visual material in highly trained musicians (Jakobson et al, 2008), and case study on music and text dyslexia (Hebert et al., 2008)[16]. An area worth exploring as a valid form of music perception, the research in music reading is still in a beginning stage comparing to that of language, and much work are expected to be done in order to make effective comparisons between the two.

III Music and Language: A Comparative View

Despite the previous discussions based on various specific areas and levels, the overall relationship between music and language has been long a subject of debate that continues to fascinate scholars one generation after another. In this section I briefly summarize and discuss the most representative overviews on this matter from recent research literature.

(1) Patel (2008): *Music, Language, and the Brain*. As previously mentioned, this book is the most comprehensive synthesis on the research literature concerning music and language up to date. As also can be seen from research summarized in this paper, Patel

[16] A cognitive deficit in either music or language reading. Patients are usually observed to be functioning normally (or even better) in other cognitive domains of music and language, such as auditory music perception and language communication.

has continuously worked on the various aspects of music and language cognition for over a decade. His important publications cover diverse areas such as rhythm, timbre and syntax from a variety of perspectives including cognitive sciences, neurosciences, acoustic analysis, and evolutionary biology. In addition, Patel has consistently showed effective analysis to the topics he dealt with, as well as effective identification of potential research topics. It should not be surprising thus that his book draws from these research literature that he already built over the years of works.

The recent interdisciplinary investigation into music and language has showed a tendency that largely differs from the previous speculative discussions and theories on the subject (e.g. Bernstein's Harvard Lecture *Unanswered Questions,* Adorno 1956, Seeger 1977), namely, the idea that "language and music cannot be considered single entities; they need to be decomposed into different component operations or levels of processing (Besson & Schon, 2001)." In light of this trend, the research (as examplified by Patel) has particularly tended to focus on the effective identification of the most meaningful points and areas of comparison, often quite different from one domain to another (e.g., the points of interesting comparisons possibly differ in sound elements perception vs. rhythmic perception between music and language). This has also resulted in the shift of the overall goal of the investigation than previously held, as Patel wrote in the Afterword of his book (Patel 2008:417): "[No matter similar or different,] Comparing music and language provides a powerful way to study the mechanisms that the mind uses to make sense out of sound."

Overall, aside from the compelling evidence that Patel have cited, his enthusiasm and belief that music and language are closely related cognitive and neural systems are apparent.[17] This is also reflected in his ability (or propelled him to possess such ability) to effectively identify the key links in each domain of comparisons between music and language as well as to propose constructive solutions, even when there are contradictory evidences that may seem to suggest the dissociation between the two in the domain in question (such as the paradoxical evidence on syntactic processing discussed earlier). Patel has also been active in broadening the scope of research by studying non-Western

[17] As I will discuss later, Jakendoff wrote in his article that even though he had pretty much the same evidence as Patel did, he would rather believe the glass is half empty rather than half full. (Jackendoff 2009)

musical systems and non-Indo-European languages, which is exactly the kind of research that needs to be done in order to bring our understanding of music and language to a more general level.

Key Links in each area of music and language, as identified by Patel (2008) include:

[1] Sound Elements (pitch and timbre): Sound category learning as a key link;

[2] Rhythm: Nonperiodic aspects of rhythm as a key link;

[3] Melody: Melodic statistics and melodic contour as key links;

[4] Syntax: Neural resources for syntactic integration as a key link;

[5] Meaning: The expression and appraisal of emotion as a key link;

[6] Evolution: Beat-based rhythm processing as a key research area.

(2) Jackendoff (2009): "Parallels and Non-Parallels between Music and Language." Although the co-author of the most influential linguistics-inspired music theory (A Generative Theory of Tonal Music), Jackendoff has constantly called for caution when drawing parallels between music and language. His recent article (2009) addresses two important questions: [1] How are language and music different? [2] In the respects that language and music are the same, are they genuinely distinct from other human activities? (Jackendoff 2009: 195). The second question, put more specifically, becomes the central argument of the article: "What cognitive capacities are shared by language and music, but not by other cognitive domain?" (ibid)

To consider this usually not addressed question, Jackendoff first outlined several general capacities shared by language and music that are also shared with other cognitive domains: [1] substantial memory capacity for storing representations (shared by the massive storage necessary for encoding the appearance of familiar objects); [2] ability to integrate stored representations combinatorially in working memory (shared with the ability to understand a complex visual environment or creating a plan for complex action); [3] creating expectations (shared with visual perception); [4] ability to achieve fine-scale voluntary control of vocal production (not required in instrumental music or sign-language, but possibly stemmed from enhanced hand control of tool making); [5] ability to imitate others' vocal production (shared with other cultural practices); [6]

ability to invent new items (shared with other cultural practices); and [7] ability to engage in jointly intended actions (shared with human widespread ability of cooperation). (Jackendoff 2009: 197)

Jackendoff also compared the different ecological functions of music and language in human life, and listed many formal similarities and differences in each of the areas of pitch, rhythm, words, syntax, prolongational structure, and complex action (See Jackendoff 2007), much of the evidence not unlike those documented by Patel (2008) and discussed in the current paper. He concluded that the two differ in many aspects of structure and function, and with the exception of the metrical grid, all aspects they share appear to be instances of more general capacities. He in turn argued "In particular, at the moment we do not have a properly laid out account of even one other capacity against which to compare language and music", and acknowledged his different view overall than Patel even though both are drawing from much of the same evidences.

At this point, it is worth considering if such approaches by Jackendoff need to be taken and how can it merit the understanding of the encoding of sound signals by the human minds. One question might worth pondering is that, in terms of shard capacity among music, language, and other cognitive domains, to what extend do they share these capacities, and what are the implications of these sharing?

(3) Pesetsky and Katz(2009): "The Identity Thesis for Language and Music." The Identity Thesis represents the contribution of formal linguistics (more specifically, generative syntax) to the study on the relationship of music and language. Different from the empirical methodology, this approach emphasizes the pure formal description of the structure of music and language on a most abstract level. Theoretically, the Identity Thesis builds on the linguistic theory of generative syntax (especially the Internal Merge theory in the analysis of wh-movement), and the GTTM in music theory developed by Lerdahl-Jackendoff. The main thesis of this theory is: "All formal difference between language and music are a consequence of differences in their fundamental building blocks (arbitrary pairings of sound and meaning in the case of language; pitch-classes and pitch-class combinations in the case of music). In all other respects, language and music

are identical (Pesetsky and Katz, 2009: 3).” In other words, it contends that the point of similarity between language and music is not their building blocks, but what they do with them.

Pesetsky and Katz first discuss the nature of GTTM, a generative grammar for music that does not resemble the appearance and form of the corresponding generative grammar in linguistics, from which GTTM received inspiration. Identifying four types (levels) of theories possible in a generative framework (type1: Analysis of particular pieces; type 2:Common properties of pieces within an idiom; type 3: Common properties of musical idioms; and type 4: Properties common to UG-M [Universal Grammar-Music] and other cognitive systems), they pointed out that GTTM did not resemble generative syntax in linguistics because these two theories belong to different types of generative theory described above. Thus, through an abstract formal analysis based on GTTM, the authors discussed the corresponding “rules” (the identical ways music and language make use of the “building blocks”) governing TSR (Time-span Reduction) and PR (Prolongational Reduction) in GTTM, and syntax and prosody in linguistics, thus “aligning” the two theories to show that their rules are much identical on an abstract level.

The Identity Thesis offers a valid explanation of the divergence between GTTM and generative linguistics, at the same time illustrates the linguistic analytical methodology that can be directly applied to similar questions in other areas of music-language comparison from a purely structural perspective. In any case, at this stage, in order to broaden the scope of this approach, it is still important to keep in mind the question that the authors themselves asked, namely, “whenever the ‘look’ of the model developed in GTTM and related literature diverges from the ‘look’ of linguistic theory… to what extent do they result from real distinctions between language and music, and to what extent do they result from non-inevitable differences in how research has been pursued in the two domains?” and vice versa.

(4) Saffran (2004): “Music and Language—a Developmental Comparison.” This overview seeks to draw research literatures from development cognitive psychology studying music and language perception at two different stages, namely, among children

and adults. The key idea in this study is that, “it seems possible that while adult musical and linguistic processes are modularized to some extent as separate entities, there maybe similar developmental underpinnings in both domains, suggesting that modularity is emergent rather than present at the beginning of life (Saffran 2004: 289).”

Appendix I

The “normalized Pairwise Variability Index” (nPVI) is defined as:

$$\text{nPVI} = \frac{100}{m-1} \times \sum_{k=1}^{m-1} \left| \frac{d_k - d_{k+1}}{\frac{d_k + d_{k+1}}{2}} \right|$$

Where m is the number of vocalic intervals in an utterance and dk is the duration of the kth interval. Two aspects of this measure make it appealing for use with music. First, the nPVI is purely relative measure of variability, i.e. the durational difference between each pair of intervals is measured relative to the average length of the pair. This normalization (which was originally introduced to control for fluctuations in speech rate) makes the nPVI a dimensionless quantity, which can be applied to both language and music. Second, the nPVI has been applied to vowels. Vowels form the core of syllables, which can in turn be compared to musical tones (i.e. in setting words to music it is quite common for each note to be assigned to one syllable). (Patel 2003b).

Bibliography

Abercrombie, D. (1967) *Elements of General Phonetics*. Chicago: Aldine.

Abraham, G. (1974) *The Tradition of Western Music*. Berkeley: University of California Press, Chap.4, 61-83.

Adorno, T.W. (1956) Music and language: a fragment. In *Quasi una Fantasia, Essays on Modern Music*, (Translated by Rodney Livingstone), London, New York: Verso.

Anvari, S.H et al. (2002) Relations Among Musical Skills, Phonological Processing, and Early Reading Ability in Pre-school Children. *Journal of Exp. Child Psychology*, 83:111-130.

Avorgbedor, Daniel. (2008) Steven Mithen, The Singing Neanderthals: The Origins of Music, Language, Mind, and Body. *Empirical Musicology Review*, 3(1): 22-35

Ayotte, J., Peretz, I., & Hyde, K. (2002). Congenital Amusia: A Group Study of Adults Afflicted with a Music-specific Disorder. Brain, 125, 238-251.

Besson, M., Friederici, A.D. (1998) Language and Music: A Comparative View. *Music Perception* 16(1):1-9 .

Besson, M, Schon, D. (2001) Comparison Between Music and Language. *Annals of the New York Academy of Sciences* 930 (June): 232-58.

Bigand, E. (2009) Editorial: Special Issue - Music and Language: 25 Years After Lerdahl & Jackendoff's GTTM. *Music Perception* 26(3): 185-186 .

Bolduc,J.(2005) Pitch Processing and Phonological Awareness. *Psychomusicology*,19(1): 3-14.

Brown,S. (2009) Review of Patel (2008), *Music, Language, and the Brain.* Musicae Scientiae, XIII(1):163-182.

_______. (2000) Musilanguage. In Wallin, N. L., Merker, B., and Brown, S. (Eds.) *The Origins of Music.* Cambridge, MA: MIT Press.

Caplan, D., Waters, G.S. (1999) Verbal working memory and sentence comprehension. *Behavioral and Brain Sciences*, 22,77-94.

Cohen, A. J. (2004) Parsing of Melody: Quantification and Testing of the Local Grouping Rules of Lerdahl and Jackendoff's "A Generative Theory of Tonal Music". *Music*

Perception 21:(4): 499-543.

Collister, B., Huron, D. (2008) Comparison of Word Intelligibility in Spoken and Sung Phrases. *Empirical Musicology Review*, 3(3):109-125.

Cook, Norman D. (2007) The Sound Symbolism of Major and Minor Harmonies. *Music Perception* 24:3, 315-319.

Cooke, Nicholas. (1994) Language, Music, & Mind by Diana Raffman. Music Perception 12:1, 156 .

Darwin, C. (1871) *The Descent of Man, and Selection in Relation to Sex.* London: John Murray.

Dauer, R.M. (1983) Stress-timing and Syllable Timing Reanalyzed. *Journal of Phonetics*, 11: 51-62.

Deutsch, Diana. (1991)The Tritone Paradox: An Influence of Language on Music Perception. *Music Perception* 8(4): 335.

Deutsch, D., Dooley, K., Henthorn, T. and Head, B. (2009) Absolute Pitch Among Students in an American Music Conservatory: Association with Tone Language Fluency. *Journal of the Acoustical Society of America*, April, 125, 2398-2403.

Deutsch,D., Henthorn, T., and Dolson, M. (2004) Speech patterns heard early in life influence later perception of the tritone paradox. *Music Perception*, 21, 357-372.

______. (2004a) Absolute Pitch, Speech, and Tone Language: Some Experiments and a Proposed Framework. *Music Perception* 21:3, 339-356.

Deustch, D., Henthorn.T., Lapidis, R. (2008) The Speech to Song Illusion. *Proceedings of 156th Meeting of the Acoustical Society of America.*

Dresher, E. (2008) Between Music and Speech: The Relationship between Gregorian and Hebrew Chant. *Toronto Working Papers in Linguistics,* 27, [np].

Forgeard, Marie. (2008) The Relation Between Music and Phonological Processing in Normal-Reading Children and Children with Dyslexia. *Music Perception* 25:4 , 383-390.

Feld, S., Fox, A. (1994) Music and Language. *Annual Review of Anthropology*, 23:25-53.

Grabe, E., Low, E.L. (2002) Durational Variability in Speech and the Rhythm Class Hypothesis. In Gussenhoven and Warner, *Laboratory Phonology*, 7. Berlin: Mouton de Gruyter.

Gibson, E. Linguistic complexity: locality of syntactic dependencies. *Cognition* 68, 1-76

(1998).

Gibson, E. The Dependency Locality Theory: A Distance-Based Theory of Linguistics Complexity. In Miyashita, Y., Marantaz, A., and O'Neil, W., eds, *Image, Language, Brain* 95-126.

Hall, R.A., Jr. (1953) Elgar and the Intonation of British English. *The Gramophone*, june 1953:6-7, repreinted in *Intonation: Selected Readings*, edited by D.Bolinger (Penguin, Harmondsworth), p282-285.

Hannon, E.(2009) Perceiving Speech Rhythm in Music: Listeners classify instrumental songs according to language of origin. *Cognition,* 111(3): 403-409.

Hébert, Sylvie. (2008) A Case Study of Music and Text Dyslexia. *Music Perception* 25:4, 369-381.

Huppertz, H.J., Knösche, T.R., Wellmer, J., Widman, G., Schulze-Bonhage, A. Sammler,D., Koelsch, S., Ball,T., Brandt, A., Elger, C.E., Friederici, A.D., Grigutsch, M., (2009) Overlap of Musical and Linguistic Syntax Processing: Intracranial ERP Evidence. *Annals of the New York Academy of Sciences* 1169:1, 494-498.

Huron, D. Ollen, J. (2003) Agogic Contrast in French and English Themes: Further Support for Patel and Daniele. Music Perception, 21:267-271.

Ilie, Gabriela. (2006) A Comparison of Acoustic Cues in Music and Speech for Three Dimensions of Affect. *Music Perception* 23:4, 319-329.

Iversen, John R. (2008) Review of "Perception and production of linguistic and musical rhythm by Korean and English middle school students" by Lydia N. Slobodian. *Empirical Musicology Review*, 3(4): 208-214.

Jackendoff, R. (2009) Parallels and Nonparallels Between Language and Music. *Music Perception* 26(3):195-204.

______. (2007) *Language, Consciousness, Culture.* Cambridge, MA: MIT Press.

Jakobson, L.S., Lewycky,S.T., Kilgour, A.R., Stoesz, B.M. (2008) Memory for Verbal and Visual Material in Highly Trained Musicians. *Music Perception* 26:1, 41-55.

Jescheniak, Jörg D. (2008) Brain Activity Patterns Suggest Prosodic Influences on Syntactic Parsing in the Comprehension of Spoken Sentences. *Music Perception* 16:1, 55-62.

Jentschke, S., Koelsch, S., Sallat, S., Friederici, A. (2008) Children with Specific Language Impairment Also Show Impairment of Music-syntactic Processing. *Journal of Cognitive Neuroscience*, 20(11):1940-1951.

Jordania, Joseph. (2006) *Who Asked the First Question: Origins of Vocal Polyphony, Human Intelligence, Language and Speech.* Tbilisi: Logos, Tbilisi State University.

Jonaitis, E. Saffran, J. (2009) Learning Harmony: The Role of Serial Statistics. *Cognitive Science* 33: 951-968.

Jones, J., Lucker, J., Zalewski, C., Brewer, C., Drayna, D. (2009) Phonological Processing in Adults with Deficits in Musical Pitch Recognition. *Journal of Communication Disorders,* 42(3):226-234.

Justus, T. Hutsler, J. (2005) Fundamental Issues in the Evolutionary Psychology of Music: Assessing Innateness and Domain-Specificity. *Music Perception*, 23: 1-27.

Kolinsky, R., Lidji, P., Peretz, I., Besson, M. & Morais, J. (2009) Processing interactions between phonology and melody: Vowels sing but consonants speak. Cognition, 112:1-20.

Kolinsky, R. (2009) Music Training Facilitates Lexical Stress Processing. *Music Perception* 26:3, 235-246 .

Lerdahl, F. (2001) *Tonal Pitch Space*. New York: Oxford University Press.

______. (2009) Genesis and Architecture of the GTTM Project. *Music Perception* 26:3, 187-194.

Lerdahl,F., Jackendoff, R. (1983) *A Generative Theory of Tonal Music*. Cambridge, MA: MIT Press.

Levitin, D., Tirovolas, A. (2009) Current Advances in the Cognitive Neuroscience of Music. *Annals of the New York Academy of Sciences* 1156:1, 211-231.

Low, E.G., Grabe, E. and Nolan, F. (2000) Quantitative Characterizations of Speech Rhythm: Syllable-timing in Singapore English. *Language & Speech*, 43(4): 377-401.

Madell, J., Herbert, S. (2008) Eye Movements and Music Reading: Where Do We Look Next? *Music Perception* 26:2, 157-170 .

Magne, C. et al (2006) Musician Children Detect Pitch Violations in Both Music and Language Better than Nonmusician Children: Behavioral and Electrophysiological Approaches. Journal of Cognitive Neuroscience, 18: 199-211.

Marin, M. (2009) Effects of Early Musical Training on Musical and Linguistic Syntactic Abilities. *Annals of the New York Academy of Sciences* 1169:1, 187-190.

McDermott, J. Hauser, M.D. (2005) The Origins of Music: Innateness, Development, and

Evolution. *Music Perception*, 23: 29-59.

Mithen, S. (2005)*The Singing Neanderthals: The Origins of Music, Language, Mind and Body.* London: Weidenfeld and Nicolson.

Nattiez, J-J. (1990) *Music and Discourse: Toward a Semiology of Music*. Princeton, NJ: Princeton University Press.

Overy, K. (2009) Dyslexia, Temporal Processing and Music: The Potential of Music as an Early Learning Aid for Dyslexic Children. *Psychology of Music*, Vol. 28, No. 2, 218-229.

Palmer, C. Hutchins, S. (2006) What is Musical Prosody? In Ross. B et al, *The Psychology of Learning and Motivation*, Vol. 46. New York: Elsevier Inc.

Parsons, K. (2006) Preserved Singing in Aphasia: A Case Study of the Efficacy of Melodic Intonation Therapy. *Music Perception* 24:1, 23-35.

Patel, A.D.(1998) "Syntactic Processing in Language and Music: Different Cognitive Operations, Similar Neural Resources." *Music Perception* 16:1, 27-42.

______. (2003) Language, music, syntax and the brain. *Nature Neuroscience,* 6:674-681.

______. (2003a) Rhythm in Language and Music: Parallels and Differences. *Annals of the New York Academy of Sciences* 999: 140-143.

______. (2006) An Empirical Method for Comparing Pitch Patterns in Spoken and Musical Melodies: A Comment on J.G.S. Pearl's "Eavesdropping with a Master: Leos Janácek and the Music of Speech." *Empirical Musicology Review*, 1(3):166-169.

______. (2006a) Musical Rhythm, Linguistic Rhythm, and Human Evolution. *Music Perception* 24(1): 99-103.

______. (2008) Music, Language, and the Brain. New York: Oxford University Press.

______. (2008a) Music and the Brain: Three Links to Language. In *Oxford Handbook of Music Psychology*. New York: Oxford University Press.

Patel, A.D., Iversen, J.R., Bregman, M.R., & Schulz, I. (2009). Experimental Evidence for Synchronization to a Musical Beat in a Nonhuman Animal. *Current Biology,* 19: 827-830.

Patel, A.D., & Iversen, J.R. (2007) The Linguistic Benefits of Musical Abilities. *Trends in Cognitive Science*, 11(9): 369-372.

Patel, A.D, Iversen, J.R, and Rosenberg, J.C. (2005) Comparing the Rhythm and Melody of Speech and Music: The Case of British English and French. *Journal of the Acoustic Society of America*, 119(5): 3034-3037.

Patel A.D. & Iversen, J.R. (2003). Acoustic and perceptual comparison of speech and drum sounds in the North Indian tabla tradition: an empirical study of sound symbolism. *Proceedings of the 15th International Congress of Phonetic Sciences, Barcelona, 2003,* pp.925-928.

Patel, A.D. & Daniele, J.R. (2003) Stress-Timed vs. Syllable-Timed Music? A Comment on Huron and Ollen (2003). *Music Perception* 21(2): 273-276.

______. (2003a). (2003b) An Empirical Comparison of Rhythm in Language and Music. *Cognition*, 87: B35-45.

Patel, A.D, Wong, M., Foxton, J., Lochy, A., & Peretz, I. (2008a) Speech Intonation Perception Deficits in Musical Tone Deafness (Congenital Amusia). *Music Perception* 25:4, 357-368.

Pearl, J. (2006) Eavesdropping with a Master: Leos Janacek and the Music of Speech. *Empirical Musicology Review,* 1:131-165.

Pesetsky, D., Katz, J. (2009) *The Identity Thesis for Language and Music*. Unpublished draft.

Pian, R.C. (1972) Text setting in the Shipi animated aria. In Laurence Berman (ed), *Words and Music: The Scholar's View; a medley of problems and solutions compiled in honor of A.Tillman Merritt by sundry hands*. Cambridge, MA: Dept.of Music, Harvard University.

Pike, K.N. (1945) *The Intonation of American English*. Ann Arbor, MI: University of Michigan Press.

Piro, J., Ortiz, C. (2009) The effect of piano lessons on the vocabulary and verbal sequencing skills of primary grade students. *Psychology of Music*, 37:325.

Ramus, F. (2002) Acoustic Correlates of Linguistic Rhythm: Perspectives. *Proceedings of Speech Prosody 2002*, *Aix-en-Provence*, 115-120. *Aix-en-Provence:* Laboratoire Parole et Langage.

Ramus, F. Nespor, M. Mehler, J. (1999) Correlates of Linguistic Rhythm in the Speech Signal. Cognition, 73: 265-292.

Roach, P. (1982) On the Distinction Between "Stress-timed" and "Syllable-timed" Languages. In D.Crystal (ed)., *Linguistic Controversies: Essays in Linguistic Theory and Practice in honour of F.R. Palmer,* 73-79. London: Edward Arnold.

Sadakata. M., Desain, P., Honing, H., Patel, A. D., & Iversen, J. R. (2004). A Cross-Cultural Study of the Rhythm in English and Japanese Popular Music. *Proceedings of the International Symposium on Musical Acousitcs (ISMA)*, 41-44. Nara.

Saffran, J.(2004) Music and Language: A Developmental Comparison. *Music Perception* 21(3):289-311.

Schellenberg, E.G., Trehub, S. (2008) Is There an Asian Advantage for Pitch Memory? *Music Perception* 25:3, 241-252.

Schön, D. (2009) Review of *Music, Language, and the Brain* by Aniruddh Patel. *Music Perception* 26(3): 287-288.

Schlaug, G., Marchina, S., Norton, A. (2008) From Singing to Speaking: Why Singing May Lead to Recovery of Expressive Language Function in Patients with Broca's Aphasia. *Music Perception* 25:4,315-323.

Seeger, Charles. (1977) *Studies in Musicology, 1935-1975*. Berkeley and Los Angeles: University of California Press.

Slobodian, Lydia N. (2008) Perception and production of linguistic and musical rhythm by Korean and English middle school students. *Empirical Musicology Review*, 3(4): 187-204.

Stegemöller, Elizabeth L.(2008) Music Training and Vocal Production of Speech and Song. *Music Perception* 25:5, 419-428.

Swinney, David. (1998) "The Processing of Discontinuous Dependencies in Language and Music". *Music Perception* 16:1, 63-78.

Thaut, Michael H.(2008) Musical Structure Facilitates Verbal Learning in Multiple Sclerosis. *Music Perception* 25:4, 325-330.

Thompson, W.F. et al. (2004) Decoding Speech Melody: Do Music Lessons Help? *Emotion*, 4: 46-64.

Ullman, M.T. (2001) A neurocognitive perspective on language: The declarative/procedual model. *Nature Reviews (Neuroscience)*, 2:717-726.

Wong, P.C.M. et al. (2007) Musical Experience Shapes Human Brainstem Encoding of Linguistic Pitch Patterns. *Natural Neuroscience*, 10:420-422.

Yung,B. (1989) *Cantonese Opera: Performance as Creative Process*. Cambridge [Cambridgeshire] ; New York : Cambridge University Press, 1989.