Language, Music, Fire, and Chess: Remarks on Music Evolution and Acquisition

John Frederick Bailyn(✉)

Department of Linguistics, Stony Brook University, New York, USA
john.bailyn@stonybrook.edu

Abstract. There is considerable debate about the evolution of both language and music cognition in human beings (see [10, 13, 24, 25] for the former, and [6, 11, 21, 33] for the latter). However, the two debates have distinct characters. In the case of language, most agree that there exists a significant biological component to the underlying cognitive system that modern humans enjoy, which in some form was either the direct or indirect product of evolutionary changes in biology. In the case of music, however, human ability in this domain has recently been compared to mastery of fire (an obvious cultural invention) [21] and specifically dismissed as not arising through evolutionary forces, understood in the standard Darwinian sense. Patel's primary arguments [21] for this relate to childhood acquisition, which he argues occurs quite differently for language and music. In particular, he claims that there is no critical period for acquisition of music perception. I examine his arguments here, coming to quite different conclusions.

1 Introduction

In this article, I first review and reject Patel's [21] arguments against music as having developed along familiar evolutionary lines, arguing that the reasoning is at best inconclusive, at worst contradictory or simply irrelevant to the core issues of the evolution of musical perception. Through the prism of well-known diagnostics, some taken from the early philosophy of mind literature ([7] a.o.), and some from Patel [21], we will see that musical perception, like basic linguistic competence, and as opposed to mastery of fire, comprises a cognitive *module*, in the sense of [8, 12], and as such is acquired developmentally much as language is. Children go through an important critical period and, if healthy with normal exposure to a local idiom, develop a steady state of musical competence in music perception of that native idiom, which requires no explicit instruction, training or study. It is as much a part of our cognitive biology as language is, and therefore just as much a product of biological evolution.

I then turn to a brief presentation of a plausible account of music evolution, based on [19], which relies on the archeological record to make the case for the evolution of music as a biological development and not a cultural invention. In conclusion, I show that some of the confusion in Patel's presentation [21] concerns the status of the mastery of musical *output* in humans, in the form of (often brilliant) facility with

P. Eismont and N. Konstantinova (Eds.): LMAC 2015, CCIS 561, pp. 27–44, 2015.
DOI: 10.1007/978-3-319-27498-0_3

instrument, voice, or with musical composition. In this area, I argue that we are dealing with something much more similar to advanced mastery of chess, a form of human accomplishment which, though fascinating in its own right, is not directly relevant to the issues of maturation and evolution of basic musical perception.

2 Patel's Basic Claim

The human ability to create and manage fire distinguishes us from all other species and appears to be universal across human cultures. However, as Patel points out [21], the universality of a human competency does not necessarily entail its biological foundation:

> The ability to make and control fire is also universal in human cultures… Yet few would dispute that the control of fire was an invention based on human ingenuity, not something that was itself a target of evolutionary forces [21: 356].

Fire is clearly a cultural invention that proved so useful that it was then taught to every succeeding generation as a matter of basic cultural knowledge. Importantly, Patel claims that the null hypothesis should be that a certain human ability not be considered part of basic human biology unless there is strong evidence to that effect:

> … the example of fire making teaches us that when we see a universal and unique human trait, we cannot simply assume that it has been a direct target of selection. In fact, from a scientific perspective it is better (because it assumes less) to take the null hypothesis that the trait in question has not been a direct target of selection. One can then ask if there is enough evidence to reject this hypothesis [21: 356].

This presupposition about which null hypothesis is more scientific could be objected to, as it considerably weakens the burden for the case against the biological nature of musical abilities. However, I will accept Patel's assumption about the null hypothesis, and show that the evidence is nevertheless strong enough to conclude about musical perception what Patel agrees we must conclude about language abilities, namely that they are unlike mastery of fire, a true cultural invention. Music, like language, has all the major hallmarks of a biological system and a mental module, under Fodor's diagnostics for the latter and under Patel's for the former. In the next section, I briefly review and apply those diagnostics to music perception.

First, however, a brief note on my assumptions about the nature of musical perception. I assume that human beings process music on a multitude of levels simultaneously (for a possible description, see *A Generative Theory of Tonal Music* [17]). These levels can include, at very least, hierarchical representations of metrical and grouping structure, tonal pitch structure and, where applicable, harmonic structure. It is important to note that Patel [21] describes musical abilities in similar terms, so that any disagreement here concerns only the evolution of these abilities, not their essentially hierarchical and representational nature.

3 Is Music Perception a Mental Module? Fodor's 1973 Diagnostics Applied

Fodor [7] provides a series of now well-known diagnostics for independent mental modules, though he does not discuss music. I will not engage in a lengthy review of the motivation for this list versus any other, but will assume that some list or other, of similar content, must be correct in identifying mental modules.

1. Characteristics of mental modules [7]

- rapidity
- automaticity
- informational encapsulation
- domain-specificity
- neural specificity
- innateness

Patel [21] does not directly address Fodor's diagnostics, though some of them clearly overlap with his own (see below). In this section I will briefly review how Fodor's diagnostics might be applied to human music perception.

3.1 Rapidity and Automaticity

It is uncontroversial that music is processed every bit as rapidly as language, under similar circumstances.[1] Experimental research clearly confirms that "this capacity rests on *fast acting* and *irrepressible processes* that enable us to extract subtle musical structures from short musical pieces" [2: 119, emphasis mine]. On this diagnostic, music processing clearly represents a unique mental module.

3.2 Informational Encapsulation

Here, Fodor has in mind the irrelevance of signals not involved in the given module's domain to perception within that domain. That is, other auditory signals, so long as they do not obscure the physical perception of the acoustic signal (by drowning it out to the ear, let us say), should not interfere with or change the organization of perceived musical signals. And indeed, what we see, smell, touch, and even hear in other domains does not affect our cognitive representation of the music. On this diagnostic, then, we have evidence of a distinct mental module at work.[2]

[1] From here on, I will assume "music" to mean familiar music within one's native musical "idiom", such as Western tonal music for presumably most, if not all, readers of this article. This is not to trivialize the significance of research into musical universals, non-tonal systems, polyphonic music and so on. See [17] for important discussion of this same assumption.

[2] Signals from other domains might, of course, combine in our *emotional* reaction to perceived music, but I contend that such interactions involve cognitive connections *across* domains, and they certainly do not affect the basic workings of the independent modules.

There is another important sense in which musical perception is informationally encapsulated. Memory, even of the identical piece of music, does not impinge on internally constructed musical expectations, in the sense of [17]. As Lerdahl [16] points out, "the unconscious processing of music continues blindly, no matter how well our conscious mind knows the music in question. To the internal processor, the musical input is always new" [16: 173].

3.3 Domain-Specificity

If there is a level of cognitive organization within a postulated module that involves principles specific to that module, not attested in other modules and not resulting from general principles of cognition, or from requirements of the cognitive interfaces, then it represents an instance of domain-specificity. The existence of entirely linguistic principles of organization, for example, is a common argument for the independence of the language module – certain purely syntactic constraints constitute such a case, and there are many more on the various linguistic levels. In musical perception, tonal pitch relations [16], constitute an equally domain-specific realm of psychological reality. Neither mathematical [9] nor psycho-physical principles can explain (even Western) tonal organization [17] in either the significance of the tonal center ("tonic") or the abstract organization of "distances" around that tone; these cognitive notions of distance do not relate to physical/acoustic distance in any measureable sense: "The general picture emerges of a theory whose … underlying constructs are constant, reflecting permanent features of musical understanding" [17: 5]. Such domain-specificity is more evidence in favor of a distinct mental module for musical representations.[3]

3.4 Neural Specificity

Can brain damage affect musical perception alone, without impacting other cognitive systems? If so, then we have an argument for neural specificity – similar to that involving aphasia in language. There is abundant evidence that such musical impairments exist. Peretz [22] and Peretz and Colheart [23] provide extensive physical evidence for "acquired selective amusia" – that is, for specific impairments that affect ONLY the musical module and in fact only certain musical sub-modules:

Patel acknowledges the existence of music-specific impairments, though he denies that this implicates natural selection, arguing that "the modularity of music processing in adults is orthogonal to the issue of selection for musical abilities. This is because modules can be a product of *development*, rather than reflecting innately specified brain specialization" [21: 357, emphasis mine]. He goes on to compare amusia with "orthographic alexia" – a reading deficit caused by brain damage. In particular, these deficits show that "there are areas in the occipitotemporal region of the left hemisphere

[3] That tonal pitch space is unique to human music in its basic abstract properties is not denied by Patel. Rather, Patel sidesteps the issue of domain specificity by arguing that modularity itself is not evidence of an evolutionary adaptation (see next section).

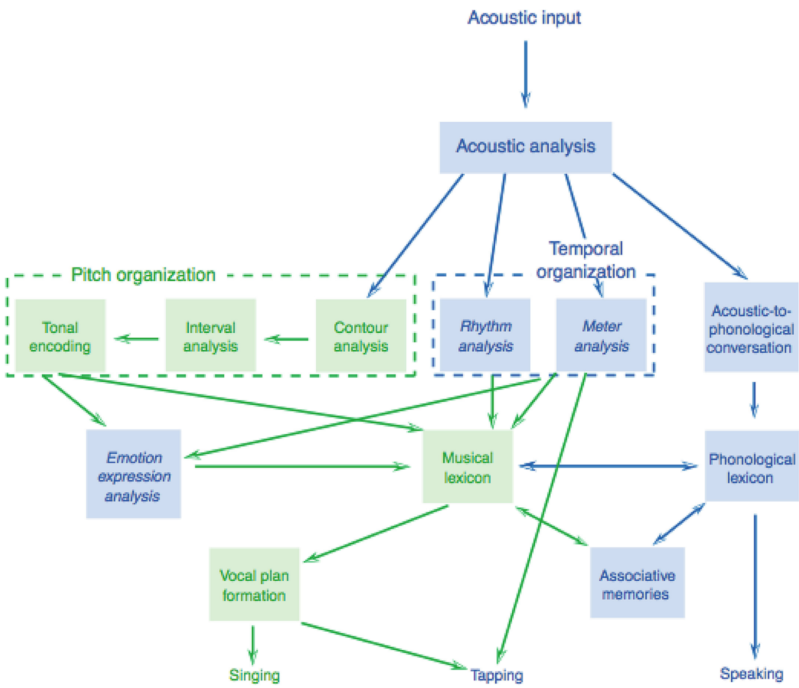

Fig. 1. Musical sub-modules showing impairments ([23] - music-specific impairments in green/light grey) (Color figure online)

that are specialized for recognizing alphabetic letters in literate individuals." Since reading is a recent human innovation, and is not universal or innate, "we can be confident that specific brain areas have not been shaped [for reading] by natural selection" [21: 357]. The argument takes the following form: if there are neurally-specific deficits attested for an ability we don't think was naturally selected for (such as alexia), we should be skeptical that we have evidence for biological evolution in other cases of neurally-specific deficits (such as amusia). However, the argument appears flawed: Patel does report findings that orthographic alexia is related to brain centers that surely *are* part of human biology, namely those involved in object recognition. Thus alexia *does* in fact reflect evidence of evolution of a neutrally specific system, one for object recognition, which is called upon in reading as well. If each of Peretz' identified cases of amusia could also be so attributed to deficits in other cognitive systems, then perhaps the argument for natural selection is weakened. But as Fig. 1 clearly shows, many of the attested deficits involve musical abilities alone (the green/light grey ones), and thus exactly resist association with other cognitive abilities. Therefore, for these cases the argument remains strong that we are dealing with evidence of musical modularity, in exactly the same sense Patel himself argues for modularity in language, and hence for natural selection.

In general, modularity arguments remain problematic for Patel's anti-evolutionary story for music. However, because his primary arguments against music as an evolved system rely on somewhat different diagnostics, I address those now separately.

3.5 Existence of Universals

The argument typically runs that for an ability to represent the product of biological evolution, we should expect it to show significant universals across idioms/dialects, that is, without significant variation. Here, then, to support the anti-evolutionary argument for music perception, we would expect Patel to deny the validity of claims of musical universals. To do so would not be unreasonable, since we have so little familiarity with distant musical idioms unrelated to our native idiom (western tonal music), and much of our music theory (including, as is obvious from its title, GTTM itself), restricts itself to western tonal music. Therefore it is difficult to substantiate claims of universals with so few idioms to compare. However, Patel tacitly acknowledges the existence of musical universals by appealing to a different strategy in this section – by casting doubt on the nature of the diagnostic – he expresses extreme skepticism about the existence of *linguistic* universals, while maintaining the biological viability of languages as a selected for system:

> grammatical universals are the focus of an interesting controversy, with some researchers arguing that these are not built into our brains but result from the convergence of a number of forces, including human limits on sequential learning, and the semiotic constraints that govern complex communication systems [21: 367]

So *linguistic* universals are in doubt for Patel. His logic appears to be as follows: If system A is biological, yet possibly shows no universals, then the presence of acknowledged universals in system B does not add to its viability as a biological system. I will leave it to the reader to decide if they find this kind of argument effective. Regardless, there are many well-known, generally accepted, musical universals, some of which are listed here (see [17] for discussion):

- All music consists of organized rhythms with organized melodies/harmonies
- Discrete perception of tones pervades all musical systems
- Melodies are perceived in terms of *motion* with regard to a tonic
- Octaves are perceived as equivalent in all musical cultures
- Octaves in all musical cultures are divided into scales consisting of 4–7 scale pitches selected from between 10 and 15 small steps

One could discuss each of these in a separate article. Taken together, though, they constitute a fairly strong case for the existence of some musical universals. Perhaps that does not seal the case in favor of natural section, but it certainly undermines the arguments against it, on this particular (and well-known) diagnostic.

4 Is Music Perception a Biological System? Patel's 2008 Diagnostics Revisited

4.1 Biological Cost of Failure

Patel argues that with any naturally selected-for system, we should expect to find a biological cost for 'failure'. That is, the giraffes that did not happen to develop long necks would have more difficulty passing their genes along, and the trait is thus selected for over evolutionary time. One unambiguously physical deficit in musical processing is tone-deafness. Patel therefore focuses his attention on that deficit, and argues as follows:

> such [musically tone-deaf] individuals appear to bear no biological cost for their deficit, and there is no evidence that they are less successful reproducers than musically gifted individuals. This is consistent with the view that natural selection has not shaped our bodies and brain for specifically musical purposes [21: 377]

However, this argument seems to entirely ignore the role of human technological manipulation of our own environment and its effect on natural selection. All that matters to an argument in favor of evolution is that there *was* a stage when people without musical perception ability *were* "less successful reproducers than musically gifted individuals" (and such a stage is presupposed by any adaptionist story, such as that of [19], discussed below). At very least, the *modern* absence of biological cost of people with such deficits (that is, absence of evidence of less successful reproduction), tells us nothing about the relevance of the trait *at the time in human evolution when it was selected for*. After all, there are humans without language, or without any of a multitude of clearly biological subsystems, who are perfectly successful reproducers in modern societies, where we have the technological ability to compensate in individual cases for quite a wide range of deficits. This is a great human achievement in many cases but clearly obscures the use of modern reproductive survival as indicative of being the product of natural selection in the original sense. I therefore contend that this argument is inconclusive and should be removed from the criteria, except to say that any adaptionist story must have such a component embedded in it to be viable, as Mithen's does (see below).

4.2 Babbling and Specialized Anatomy

The vocal tract is well adapted for both music and speech. Furthermore, children go through a babbling period for both music and language. Patel acknowledges that

> babbling… and the anatomy of the vocal tract could all reflect adaptations for an acoustic communication system that originally supported *both* language and vocal music. *It is … ambiguous which domain (music or language) provided the relevant selective pressures for these features of human biology*" [21: 371-2, emphasis mine]

This is a remarkable statement, if the purpose is to argue *against* adaptive pressures for music, since it simply says that it is indeterminate whether musical or linguistic adaptive pressures (if the two can be distinguished) underlie the existence of a babbling stage of acquisition and of the development of specialized anatomy. If we cannot tell what the source of the selective pressure was, then these features of human biology simply cannot be used to differentiate music from language in terms of evolution! (Not to mention that it is being admitted at the outset that music *can*, in principle, provide selective pressures). At very least, these diagnostics do not speak *against* a possible evolutionary story for musical processing abilities (and possibly speak in favor of one). Again, at best (for Patel) inconclusive.

In the next sections I turn to Patel's central arguments against natural selection of music– that humans do not show a robust predisposition for music and do not go through a critical period for music acquisition as we do for language.

4.3 Predisposition and Precocious Learning

One of the points of contention here is the speed and regularity of the early childhood acquisition process. The implicit comparison is with language, where it is generally accepted that there is a strong predisposition and a robust critical period ([3, 4, 6, 27], a.o.). Patel [21] reports studies that show relatively slow acquisition of the musical notion of key membership, a central piece of knowledge of tonality in western tonal music. Primarily, this is based on his reading of Trehub and Trainor's studies [29, 30] comparing infant vs adult perception of in-key vs out-of-key changes in tones. The experiments work as follows: infants and older children are presented with tonal changes, some of which maintain key membership and some of which do not. Various behavioral measures determine the extent to which the children are more aware of the in-key changes than the not-in-key changes. Patel summarizes: "[there is] evidence that implicit knowledge of key membership is not in place by 8 months of age… [but] somewhere [before] 5 years, children develop a sense of key membership…" [21: 372]. He then reasons as follows:

> if musical pitch abilities had been the target of natural selection, one would expect accurate perception and production of musical pitch patterns to be learned far more quickly [than these studies show]. The slow development is especially striking given that *the music children hear* (e.g. nursery songs) *is especially strong in its tonal structure*, with few out-of-key notes [21: 372, emphasis mine]

First, a comment on the last point (that the nature of nursery songs should influence the speed of acquisition). It is perfectly possible that nursery songs share analogous characteristics with *motherese*, the exaggerated intonational patterns and lexical simplification used in child-directed speech, but there is no evidence that these impact the *speed* of the acquisition process with language [20]. It is also inaccurate to claim these songs have "stronger tonal structure" than other input music – that's akin to saying certain linguistic input is "more grammatical" than other input. We know that complex linguistic structures are simply not perceived at certain early stages (see [18] as well as [26] for discussion of "zero level" triggers accessing main clause information only in setting linguistic parameters). Presumably the same would apply to more complex musical structures –basic relations would be attended to, more complex ones not. At very least, the existence of simplified infant-directed song does not in any theory of acquisition entail that things should be different than with language, where we also find infant-directed speech.

More important, however, is Patel's claim that the music acquisition process is, in fact, relatively slow, in developmental terms. In actuality, the observed rate for acquisition of key membership looks quite similar to the trajectory for learning certain linguistic distinctions *in one's native language*, which (western) tonal relations are analogous to (specific tonality systems are of course not universal). That is, we know that although tonality as a cognitive notion may well be universal (see above), its particular instantiation has to be learned, like one's native lexicon in language, a process that continues throughout childhood and beyond, or like complex syntactic relationships, some of which do not fall into place until late in the acquisition period, such as the development of A-chains, which have been shown not to be correctly handled until past age 8 ([1], see also [26]). So if key membership is in place by age 5, then mastery of one's native musical idiom, in all it idiosyncrasies, does not seem to be any slower in acquisition terms than many analogous aspects of linguistic competence.

Crucially, in the studies reported by Patel, 8 month old infants "detected both kinds of changes" (that is, both in- and out-of-key changes). This shows exactly a strong early predisposition to one of the central aspects of musical perception – differences in scalar tones. One could imagine the distinctions not even being attended to at all until a much later age – this might well be what we expect if music were simply a cultural invention, as Patel claims.

Patel also reports on foot-tapping studies (where subjects are asked to tap their feet to the metrical beat of a piece of music) in which non-musicians fare worse than musicians, and claims that "humans appear far more uniform in their linguistic than in their musical abilities. Although some normal people are certainly more fluent speakers or have keener ears for speech than do others, these variations seem minor compared to the range of musical abilities in normal people" [21: 375]. It is entirely unclear on what basis Patel is able to draw this conclusion, other than the obvious distinctions between trained musicians and ordinary healthy music perceivers – clearly the former can have abilities the latter do not have, such as being able to read music, play an instrument, analyze musical relationships, compose music and so on. With language, however, one could also identity a huge range of differences among humans, if we included in the picture literacy (or even degrees of literacy), public speaking ability, the kind of creativity expressed through literature, poetry and so on. Patel is confounding musical *creativity* and *output* skills with the basic human musical competence of Lerdahl and Jackendoff's "experienced listeners" where there seems to be very little variation. The fact that trained musicians perform better on foot-tapping experiments than non-trained individuals is beside the point – it shows merely that training helps one succeed at such exercises just as explicit linguistic training would help one explicitly identify, say, syllable boundaries in one's native language. But even those who cannot explicitly identify syllable boundaries, or even know what a syllable is, still apply internalized phonological rules requiring knowledge of syllable boundaries in exactly the same way as trained linguists. Analogously, all listeners perceive the metrical structure of a piece equally accurately, and create the same representations for it, even if some cannot perform output tasks such as foot-tapping as accurately.

The overall conclusion is that the studies cited by Patel do not in fact show that acquisition of basic music processing abilities proceeds more slowly than language acquisition does. Nor does Patel present evidence for lack of strong predisposition with regard to music. This, combined with the evidence summarized in [31], leaves us with a comparable situation with music and language acquisition. Critical period effects, which I turn to next, reinforce this conclusion.

4.4 Critical Period

For Patel, the existence of a critical period in acquisition of a cognitive function (such as language or music), is practically synonymous with it having undergone a process of natural selection. Recall that his null hypothesis is that any ability should be considered a cultural invention unless and until there is strong empirical evidence to the contrary, which for language he claims there is, in particular with regard to a critical period. "Until good evidence appears showing rapid development of musical skills *that are not related to language* or to general principles of auditory function, there is no reason to reject the null hypothesis that music has not been a target of natural selection" [21: 374].

This is not an unreasonable way to proceed. By most accounts, however, this leads to the exact opposite conclusion from what Patel proposes. Specialists in child acquisition of basic music processing abilities (again, not to be confused with musical output skills), agree that a critical period is clearly present (see [22, 31, 28] among many others).

> studies of genetics, behavior, and brain structure and function in conjunction with the experiences of auditory deprivation and musical enrichment, … conclude that *there is more supporting evidence for critical periods for basic than for more complex aspects of musical pitch acquisition…* [28: 262-4]

Remarkably, Patel does not directly engage with the existing literature on musical acquisition. Rather, he hypothesizes that experimental results can be ignored in the face of a larger theoretical point. Consider carefully the following selection from [21] (for ease of reference, I have numbered the assertions here):

(1) *even without doing … experiments, it seems that even if a critical period effect for music is found, the effect will be rather weak compared to language.*
(2) One reason to suspect this is that some highly accomplished musicians did not start playing their instrument until after the age of 10. (To take just one example, George Gershwin was introduced to the piano at 13.)

First, with regard to (1), clearly we do not want to promote an anti-experimentalist approach to the issues at hand. The questions are empirical, and there is extensive experimental work reported in the literature, though clearly much more remains to be done. But to argue that that "even if a critical period effect for music is found, the effect will be rather weak compared to language" is to presuppose an empirical result on the basis of ones biases, rather than to trust the scientific method. This clearly contradicts Patel's general stance on experimental work, evident in successful discussion throughout the rest of his important book, and must be viewed with extreme skepticism, unless the theoretical argument against the hypothetical experimental results is overwhelming. However, as (2) and (3) show, this is far from the case.

(2) implies that we are discussing the existence of a critical period for *trained* musicians, rather than for typical human 'experienced listeners' without musical training.[4] However, that is not the appropriate subject matter for discussion of a critical

[4] This is not to deny the importance of understanding technical musical skills, musical creativity, output abilities and so on. I return to those issues in the conclusion.

period for basic music processing abilities. The issue is not the status of trained musicians, or anything about "output", but rather the status of basic human competence. To be clear: "the experienced listener" is *not* a trained musician, s/he is simply a healthy human being who was exposed to some kind of musical input as a child and who processes musical signals into complex and uniform representations, about which s/he has strong intuitive knowledge.

Naturally, we might wonder about experimental results comparing the brains of trained musicians vs. those with nothing beyond basic musical exposure. Here, again, one must listen to the experts:

> Given the current stage of research, the differences found between the brains of musicians and nonmusicians remain rather weak in light of the considerable difference in musical training that exists between the two groups. … To our view, *these differences are negligible in front of the large overlap in brain activities found in musically trained and untrained listeners* [2: 124, emphasis mine].

That is, training is irrelevant to the basic cognitive ability in question.

The example of George Gershwin (3), is misleading in the same way. Gershwin can be studied as an example of musical genius, a brilliant composer, and wonderful cultural producer. And yes, his formal musical training began after the end of the traditionally understood critical period. But this is beside the point. If Gershwin had not been exposed to any music at all *of any kind* before age 13, then surely his successes (or indeed any normal human's success in acquiring the complex system of processing one's native musical idiom) could be taken as a problematic case for the critical period hypothesis. But his basic cognitive musical processing abilities were surely in place, just as anyone else's are, when his musical training began. If by musical abilities one means ability to perform, produce, compose and analyze music, then of course the question of a critical period changes to one of a possible critical period for this and many other aspects of human creativity, genius, acquired technical skill and so on. Much of Patel's reasoning confounds the two distinct kinds of musical ability, and could in the end be attributed to a kind of terminological confusion, rather than a deep misunderstanding of the importance of a critical period for acquisition of basic musical processing ability. Experimental work needs to therefore be consulted, and on this the Trainor/Trehub results seem quite clear: basic human music processing abilities are subject to a clear critical period.

4.5 Existence of Alternate Modalities

Patel appeals to the fact that language emerges in distinct modalities – it is not only spoken, but also signed, and sign-language research has clearly shown us that the level of complexity is similar across the modalities. He points to the fact that there does not seem to exist anything like "signed music" which could utilize a visual rather than auditory modality but which shares the same mental representations as usual music. This is an interesting argument that requires careful attention. Cognitive studies of dance, for example, may prove an interesting realm in which to search for alternate musical modalities. However, because basic musical cognition is a form of processing (just as visual processing does not have an output component), the issues are difficult to evaluate. At this stage, I do not find anything conclusive in the existence of both aural and visual modalities for languages that suggests that music is merely a cultural invention. I leave the empirical issues here to further research.

4.6 Summary

In Sect. 3, we found strong evidence for the independent modularity of music, in the sense of Fodor [7] and Jackendoff [12], using 5 well-known characteristics, only one of which Patel makes different claims about (#5 below). In Sect. 4, we looked at 7 of Patel's own diagnostics for music vs language, which start from his comparison of music to fire, rather than language. For 3 of his 7 diagnostics (#s 9, 10, 11), we have seen strong reason to doubt Patel's claims that music and language differ in any significant way (predisposition, robustness of acquisition and the critical period). For the other 4 properties, Patel in fact concedes music has the same properties as language, and that either language or music could have provided the selectional pressures to create the current situation. The following chart summarizes the findings (making some uncontroversial assumptions about the nature of fire as a 'cognitive' system) (Fig. 2.)

	Language	Fire	Music (P)	Music (JFB)
0. (Universality)	✓	✓	✓	✓
• Modularity:				
1. Rapidity and automaticity	✓	*	✓	✓
2. Informational encapsulation	✓	*	✓	✓
3. Domain specificity	✓	*	✓	✓
4. Neural specificity (biological defects)	✓	*	✓	✓
5. Existence of universals	✓/*	*	*	✓
• Predisposition:				
6. Biological cost of failure to acquire	✓	??	*	NA
7. Babbling	✓	*	✓	✓
8. Specialized anatomy	✓	*	✓	✓
9. Predisposition	✓	*	*	✓
10. Precocious learning	✓	*	*	✓
11. A critical period for acquisition	✓	*	*	✓
12. Existence of alternate modalities	✓	*	✓	??

Fig. 2. Summary of 5 of Fodor's [7] and 7 of Patel's [21] diagnostics for language, fire and music

Nothing we have seen supports Patel's assertions of a radical difference between music and language.[5] If this article is on the right track about the four diagnostics shaded dark grey above, where I differ from Patel, then we find no significant distinctions between language and music, either in the modularity literature, or based on the diagnostics taken from Patel's own chapter on music evolution. Patel's conclusions about music evolution are all the more paradoxical given his own work, and that of [15] on shared linguistic and musical resources (the "shared syntactic integration resource hypothesis" - SSIRH) also presented in [21] (mostly in Chap. 5; the evolution claims are in Chap. 7). "Cross-domain interference effects show that although the two domains have distinct syntactic representations (e.g., chords vs. words), *they share neural resources for … integrating these representations during … processing"* [21]. The SSIRH specifically claims a common resource center in the brain for language and music, something unlikely, though not impossible, to be found if one of the abilities were a biological development and the other a cultural invention. This, along with the diagnostics above that were rejected as being indistinguishable between linguistic and musical pressures on evolution (such as development of human vocal tract anatomy) compel us to take seriously stories of music/language co-evolution. A co-evolution story allows early human *musical properties* to be associated with *communicative and expressive functions*, without implicating speech itself. In the next section I turn to one such story.

5 A Plausible Story of Music Evolution

As noted above, Patel's work on shared music/language resources, the common development of various physical traits that support both language and musical abilities, as well as a host of theoretical work showing a strong parallelism between the combinatoric system of both modules (esp. [14]) support the "musi-language" model of language evolution schematized below, from [33] (Fig. 3):

It is not the purpose of this article to promote a particular view of music evolution – I have not studied the complex issues of language or music evolution enough to be qualified to do so. However, because I am convinced as a cognitive scientist that the claim that music is purely a cultural invention cannot be maintained, it is worthwhile in this context to point to a plausible evolution story, one that is consistent with what we know about both the similarities and the differences between music and language. For that purpose, I turn to Mithen [19].

5.1 Mithen's HMMMMM

Mithen [19] argues that early humans had a common vocalization system with elements of both music and language, which he calls "Hmmmmm" (Holistic, manipulative,

[5] Perhaps his claim stems from the difficult task of determining what adaptionist pressure might have helped music develop in human evolution. However, Patel does not enter into the debate over language evolution, and to conclude that some sort of evolutionary pressures were involved in the development of human musical processing ability would not commit him to a particular story.

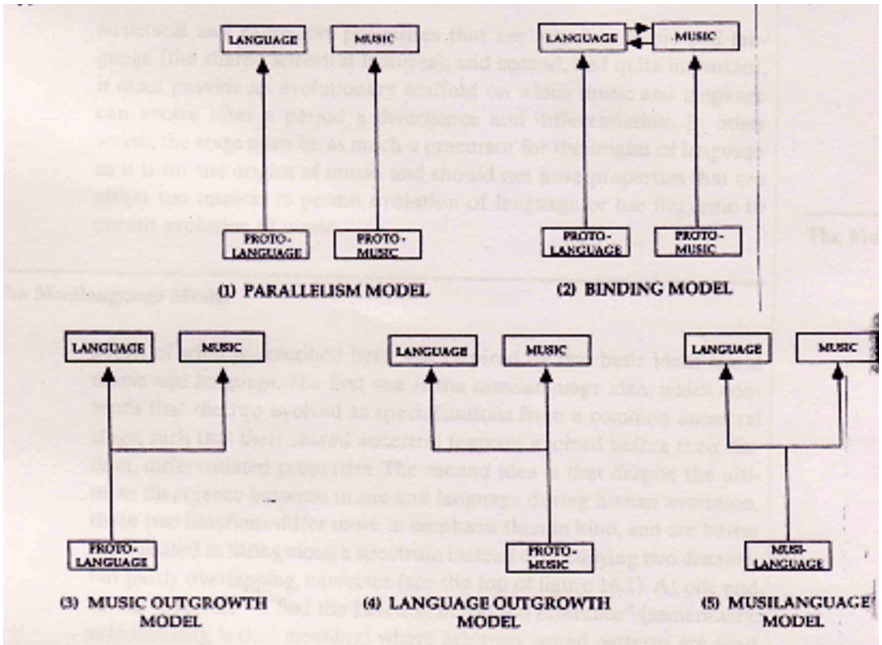

Fig. 3. Possible language and music evolution stories [33]. The musi-language model that best describes Mithen's proposal [19] is shown in the bottom right of the figure.

multi-modal, musical, mimetic). The Hmmmmm system, active around 2 million years ago, is fully consistent with the archeological record, and what is known about human cognition and group interaction at that time. Its primary function was connected with social cooperation and establishment of trust within groups, allowing longer term group-beneficial decisions to prevail over shorter-term individualistic goals. The time-frame is given in Fig. 4 on the next page.

The Hmmmmm system had the following characteristics:

- **Each utterance was *holistic* - it carried a single meaning** - there was no semantic compositionality, explaining why there was little or no change from 3 M to .75 M years ago. (They had fire and tools, but no fireplaces)
- **Each utterance was *manipulative*** - it was an attempt to achieve a resulting interaction or reaction; appropriate to the level of theory of mind known at the time
- **Each utterance was *multi-modal*** - it contained rhythm, pitch, tone, melodic contour, accompanying gesture, etc.
- **Each utterance was *musical*** - the elements listed under multi-modality comprise the basic building blocks of all human music
- **Each utterance was *mimetic*** - it imitated or mimicked (almost, but not quite 'referred to') something in the surrounding environment, be it hunting parties, food opportunities, social interactions, emotional expression, group bonding, sexual attraction and so on

The Hmmmm story is consistent with the most successful adaptive narrative of music evolution: as a way to 'coordinate coalitions' (group activity). Mithen describes it as follows:

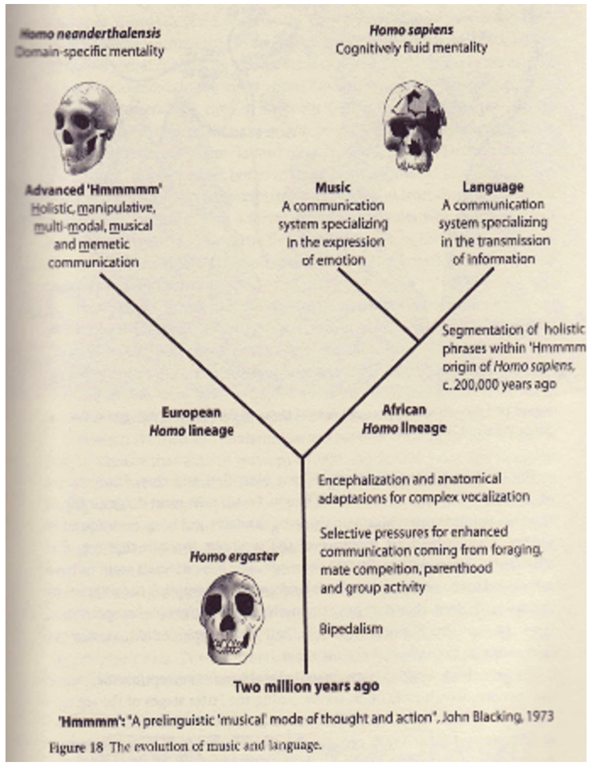

Fig. 4. Mithen's [19] picture of the evolution of language and music

> Hominids would have frequently and meticulously examined the likely intentions, beliefs, desires and feelings of others members of a group before deciding whether to cooperate with them. But on other occasions simply trusting them would have been more effective, especially if quick decisions were necessary. As a consequence, those individuals who suppressed their own self-identity and instead forged a group identity by shared "Hmmmmm" vocalizations and movements, with high emotional and hence musical content, would have prospered.
>
> Joint music-making served to facilitate cooperative behaviour by advertising one's willingness to cooperate, and by creating shared emotional states leading to "boundary loss"/"we-ness"/"coupling"/"in-group bias" With the evolution of *Homo ergaster* and full bipedalism, 'Hmmmmm' gained additional musical qualities, while further selective pressures for its evolution arose from the need to transmit information about the natural world, to compete for mates and to care for infants [19: 217-8].

Later, language and music divided, each becoming associated with modern human abilities, as we know them, while retaining certain cognitive points of similarity. Thus the shared resources of [21] and the syntactic identity identified in [14] find a natural source – their common history (claimed, I should certainly point out, by none of them).

Cognitive archeology is a growing field, and we can one day expect to understand far better than we do now how human cognitive evolution proceeded, especially once we have a better understanding of its neural instantiation. But that day may be some time off, and we owe it to those working then to be sure we are now asking the right questions, and looking in the right places for advancement in this area. The musilanguage approach, and a strong understanding of the biological nature of musical processing abilities in modern humans, are the proper starting point.

6 Conclusion: Thoughts on Creativity, Genius, and Technical Mastery

Of course it remains a fascinating question as to what is different about brilliant musicians, especially creative geniuses and dazzling instrumentalists such as Mozart or Bernstein or Prokofiev or Janis Joplin or Montserrat Caballé or Lang Lang. To be absolutely clear; I am not claiming that natural selection is involved in the development of musical genius of this kind. (Nor is Patel, despite the confusing discussion of Goerge Gershwin as having only begun formal musical training at age 13.) However, some of the experimental work comparing brain reactions of musicians vs non-musicians bears this character, presupposing that all human musical abilities are a reduced form of musical creativity and output skills. I maintain on the contrary that it is imperative to distinguish basic cognitive abilities that all (healthy) humans share from learned (or innately enhanced) abilities that we find in a small minority of gifted individuals. Surely the question of how to understand musical genius is a subset of a larger questions about human creativity, and human "genius", which are so poorly understood by cognitive scientists (though see [32] for a recent attempt).

Here, we are at a loss similar to what we encounter in trying to describe what is different about the mind of chess masters. A game invented for human entertainment now captivates the world of cognitive and computer science because of its intense complexity, and the facility with which some human minds grasp it to a degree that cannot be matched by computers, by description, by discussion etc. Little is known about genius of this sort, except that it is, in addition to everything else, the product of intense study and hard work (even for those most gifted). Gobet and Campitelli [8], for example, found a strong correlation between the number of hours chess players have dedicated to chess (deliberate practice) and their current rating:

i. Unrated players reported an average of 8,303 h of dedication to chess; rated players reported 11,715 h; Fide Masters reported 19,618 h and International Masters reported 27,929 h
ii. Stronger players tend to own more chess books (and read them) than weaker players. As an individual activity, *reading chess books is the most important predictor of chess skill.*

> So what does it take to become an expert-a chess master, a concert pianist, or a tennis champion? Studies of expertise … show some consistent parallels across domains. One needs a certain amount of time; …. In music, for instance, *reaching a professional level of expertise appears to take l0.000 h*, with a further 25,000 h of added exposure to music-related activities (listening to pieces, reading scores, taking relevant classes, etc.), (see Ericsson 1996). Similar times hold for gaining expertise in complex games like chess … [5, emphasis mine]

Of course if the training takes place at a young age, the results may be more impressive; that depends on the nature of a critical period for learning in general. At very least we know that no study at all is required to achieve complex mastery of basic musical processing, and that the human mind is endowed with the unique modular ability to organize musical input into complex hierarchical representations.

References

1. Babyonyshev, M., Ganger, J., Pesetsky, D., Wexler, K.: The maturation of grammatical principles: evidence from Russian unaccusatives. J. Linguist. Inq. **32**(1), 1–44 (2001)
2. Bigand, E., Poulin-Charronnat, B.: Are we "experienced listeners"? A review of the musical capacities that do not depend on formal musical training. J. Cogn. **100**, 100–130 (2006)
3. Brown, R., Hanlon, C.: Derivational complexity and order of acquisition in child speech. In: Hayes, J.R. (ed.) Cognition and the Development of Language, pp. 155–207. Wiley, New York (1970)
4. Crain, S., Lillo-Martin, D.: Linguistic Theory and Language Acquisition. Blackwell, Cambridge (1999)
5. Clark, E.: Critical Periods, Time, and Practice. U Penn Working Papers in Linguistics, vol. 9.2 (2003)
6. Fitch, W.T.: The biology and evolution of music: a comparative perspective. J. Cogn. **100**, 173–215 (2006)
7. Fodor, J.: Modularity of Mind: An Essay on Faculty Psychology. MIT Press, Cambridge (1983)
8. Gobet, F., Campitelli, G.: The role of domain-specific practice, handedness, and starting age in chess. J. Dev. Psychol. **43**(1), 159–172 (2007)
9. Harkleroad, L.: The Math Behind the Music. Cambridge University Press, Cambridge (2006)
10. Hauser, M.D., Chomsky, N., Fitch, T.: The faculty of language: what is it, who has it, and how did it evolve? J. Sci. **298**, 1569–1579 (2002)
11. Huron, D.: Is music an evolutionary adaptation? In: Wallin, N.L., et al. (eds.) The Origins of Music, pp. 57–75. MIT Press, Cambridge (2001)
12. Jackendoff, R.: Consciousness and the Computational Mind. MIT Press, Cambridge (1987)
13. Jackendoff, R.: Foundations of Language. Oxford University Press, New York (2002)
14. Katz, J., Pesetsky, D.: The Identity Thesis for Music and Language. LingBuzz (2011)
15. Koelsch, S., Gunter, T.C., Wittfoth, M., Sammler, D.: Interaction between syntax processing in language and in music: an ERP study. J. Cogn. Neurosci. **17**(10), 1565–1577 (2005)
16. Lerdahl, F.: Tonal Pitch Space. Oxford University Press, Oxford (2001)
17. Lerdahl, F., Jackendoff, R.: A Generative Theory of Tonal Music. MIT Press, Cambridge (1983)
18. Lightfoot, D.: Why UG needs a learning theory: triggering verb movement. In: Jones, C. (ed.) Historical Linguistics: Problems and Perspectives, pp. 190–214. Longman, London (1993)
19. Mithen, S.: The Singing Neanderthals: The Origins of Music, Language, Mind and Body. Weidenfeld & Nicolson, London (2006)
20. Nelson, D.G.K., Hirsh-Pasek, K., Jusczyk, P.W., Cassidy, K.W.: How the prosodic cues in motherese might assist language learning. J. Child Lang. **16**(01), 55–68 (1989)
21. Patel, A.D.: Language, Music and the Brain. Oxford University Press, Oxford (2008)

22. Peretz, I.: The nature of music from a biological perspective. J. Cogn. **100**, 1–32 (2006)
23. Peretz, I., Coltheart, M.: Modularity of music processing. J. Nat. Neurosci. **6**(7), 688–691 (2003)
24. Pinker, S., Bloom, P.: Natural language and natural selection. J. Beha. Brain Sci. **13**, 707–784 (1990)
25. Pinker, S., Jackendoff, R.: The faculty of language: what's special about it? J. Cogn. **95**, 201–236 (2005)
26. Snyder, W.: Child Language: The Parametric Approach. Oxford University Press, Oxford (2007)
27. Snyder, W., Lillo-Martin, D.: Principles and parameters theory and language acquisition. In: Hogan, P. (ed.) The Cambridge Encyclopedia of Language Sciences. Cambridge University Press, Cambbridge (2011)
28. Trainor, L.J.: Are there critical periods for music development? J. Dev. Psychobiol. **46**, 262–278 (2005)
29. Trainor, L.J., Trehub, S.E.: A comparison of infants' and adults' sensitivity to western musical structure. J. Exp. Psychol. Hum. Percept. Perform. **18**, 394–402 (1992)
30. Trainor, L.J., Trehub, S.E.: Key membership and implied harmony in western tonal music: developmental perspectives. J. Percept. Psychophys. **56**, 125–132 (1994)
31. Trehub, S.E.: Musical predispositions in infancy: An update. In: Peretz, I., Zatorre, R. (eds.) The Cognitive Neuroscience of Music, pp. 3–20. Oxford University Press, New York (2003)
32. Turner, M.: The Artful Mind: Cognitive Science and the Riddle of Human Creativity. Oxford University Press, Oxford (2006)
33. Wallin, N.L., Merker, B., Brown, S. (eds.): The Origins of Music. MIT Press, Cambridge (2001)