Musicæ Scientiæ
Discussion Forum 5, 2010, 231-255

On the non-existence of music: Why music theory is a figment of the imagination

GERAINT A. WIGGINS, DANIEL MÜLLENSIEFEN AND MARCUS T. PEARCE
Centre for Cognition, Computation and Culture
Goldsmiths, University of London

• **Abstract**

We argue for an approach to the theory of music which starts from the position that music is primarily a construct of human minds (and secondarily a social construct) and contrast it with the approach implicit in the work of some music theorists, which treats music as though it were an externally defined quasi-Platonic absolute. We argue that a natural conclusion of this approach is that music theory, while already being a kind of folk psychology, can benefit from being more explicitly informed by music cognition studies. We give examples from work in the computational modelling of music cognition, following our approach, which attempts to place each musical phenomenon in an ecological context motivated by evolutionary considerations, and which aims to explain musical phenomena independently of the explicit intervention of the theorist. We argue that only thus can a theory be said veridically to explicate the phenomenology of music. We place our argument in context of the Generative Theory of Tonal Music (Lerdahl & Jackendoff, 1983), Generative Linguistics, and other papers in the current volume, and compare them all with results of modelling studies based on our espoused approach.

Keywords: modelling, syntax, semantics, music and language.

Introduction

This paper is intended as a synoptic response to the other papers in this volume, placed in the context of the Generative Theory of Tonal Music (Lerdahl & Jackendoff, 1983). In such a rich and diverse collection of work, it is well nigh impossible to draw all authors together in one common narrative, and so we do not attempt to do so here — in any case, the authors have presented their own work, and there is nothing to be gained in a summary from us. Rather, we take a stance which might be thought of as challenging to the starting point of Lerdahl and Jackendoff

(1983) and certainly to the overall practice of attempting to capture, in Chomskian grammatical formalisms, the structure perceived by human listeners in music.

Our starting point is that there is *no such thing* as Music, except as created, perceived, and experienced in the human mind. In other words,

Music, in its own right, does not exist.

This claim may be seen as an outrageous falsehood, or as fatuously self-evident, depending on the reader's position; and some readers will see it merely as sophistry. In the current paper, we will argue that it is philosophically useful because it forces us to question assumptions we may make at the very beginning of our research, which have profound consequences for what we do and write about its outcomes.

We begin by placing our argument in a philosophical context, from which we develop our own starting point in contrast to the Chomskian-influenced, quasi-Platonic approaches evident in this volume. Along the way we discuss various points arising from the specific contrasts offered by the current papers, some of which support our position. Finally, we loosely contrast our own IDyOM model with GTTM, highlighting how different starting assumptions naturally lead to different kinds of theory.

Where and what is music?

The word "music" is used in many different ways. Some of them are analogical, and not relevant here, but nevertheless there are sufficient different usages for meanings to be confused. For example, a band member may ask her colleague to pass the "music", meaning the score or instrumental part, from which she is to play; engineers sometimes report on "music" processing, meaning the manipulation of audio signals which are generated by musical performance; teenagers are proud of their "music" collection, which is actually a quantity of CDs, tapes or MP3 files; and some academics discuss "music" analysis, meaning (among other things) the prediction and attempted explication of the effect of performances of scores on listeners, or the reconstruction of compositional principles and strategies applied by a composer in a specific work.

Babbitt (1965) proposes a view of psychological music representations divided into categories based on the kind of external representation they are derived from: the *acoustic* (or physical), the *auditory* (or perceived) or the *graphemic* (or notated) *domain* (in which we must, in the current times, include representations such as the digital encoding used in CDs and mp3 files). None of these is presented as the definitive "Music". We follow this view here, taking the philosophical stance that the mysterious thing that is Music is actually something abstract and intangible, and which does not have real existence in itself, but which is *described* by all three of these representational domains; in a sense, something like the notion of the Platonic ideal, as expressed, for example, in *The Republic*, but *without actually existing in the real world*[1]. It is therefore reasonable to think of all of these domain-specific representations as being *aspects* of Music, but none of them *is* Music, individually.

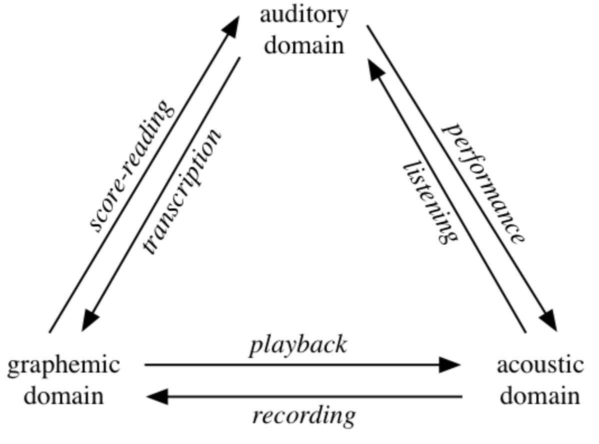


Figure 1.

Illustration of Babbitt's domains, with our addition of transformations between them, quoted from Wiggins (2009).

Figure 1 illustrates the relationships between the domains; Music, itself, lies somewhere in the middle, referred to and described by all the domains, but not actually *being* any of them. Further, even though the diagram does not include a temporal dimension, it is evident that one is needed, both within a musical experience and also along a historical timeline. In quasi-mathematical terms, it is perhaps helpful to think of each of the domains, taken diachronically, as a projection of the "whole Musical object", which is itself not directly available.

While it may not be possible to specify exactly what Music is, since we can only pin it down by saying "all these things describe it", we can perhaps identify *where* it is. In recent years, the audio engineering community has coined the term "the semantic gap", to refer to musical information which stubbornly refuses to be extracted from audio signals in isolation; the same community has identified what it

(1) Note that this is a very important difference. We will argue later that the assumption that music is a real-world object — in the sense of Popper's World I — and is therefore amenable to study, in the same way as physical objects, under the standard scientific method is at the root of many misconceptions.

calls a "glass ceiling" at about 70% accuracy in the results it can achieve by audio processing alone (*e.g.*, Aucouturier & Pachet, 2004). Most musicians and psychologists would probably agree that this gap is filled by what a listener, performer or other musician brings to the musical experience.

More than this, however, music can exist entirely *without* audio, in the minds of skilled listeners, and, albeit exceptionally, score readers. Musical imagery studies show that brains are activated, when people listen to music, in very similar ways to when they imagine it (*e.g.*, Zatorre & Halpern, 2005; Kraemer *et al.*, 2005).

The experience of the audio engineers, then, confirms that the audio signal is less than the whole, and imagination of skilled musicians shows audio stimulus to be unnecessary; and since most people cannot hear a score, but can listen to music perfectly well, Music cannot primarily reside in notation. Since, we suggest above, the difference is in the listener, performer or other musician, we can only argue that the place where Babbitt's three domains come together is really in the human mind/brain. It is therefore perhaps useful to make some rather sharper distinctions in terminology than is common: what happens on the outside of an ear might strictly be called "acoustic"; what happens on paper, or as notation carried by other media, including the "representations" in computer science (Wiggins *et al.*, 1993), might be described as "graphemic"; and the only place left for the rest can only be firmly in Babbitt's auditory domain — that is, the psychological effect of a stimulus, audio or notated, which is the human response to Music. It is still worth emphasising that there is a larger meaning of "Music" which arises from the combination of all the domains in their diachronic sociological context; but the fundamental *source* of Music — that without which Music cannot exist — is *in the mind*. In this view of the musical world, both the audio signal and the notation are stimulus and/or result, applied to or derived from a cognitive process, depending on what activity is taking place. Wiggins (2009) gives a fuller version of this argument, from the engineering perspective.

Our starting position, then, is that Music is *fundamentally* a psychological entity, which leaves traces in the real world (audio signals and notation) which are themselves musical stimuli — much as light cannot itself be seen, but leaves traces everywhere around us as it rebounds from objects. This position, in turn, raises questions about the study of music. In particular, is it adequate to propose a "theory of music" which is based mostly, if not exclusively, on the study of the traces left by music in the world, rather like studying the dynamics of ice floes by looking only above the water-line? We address this question in the next section.

What is a theory of music?

To begin with, it is important to acknowledge the difference between the meanings of the word "theory" as applied in "music theory" and, on the other hand, in "scientific theory".

According to Popper (1959) and Lakatos (1978), a scientific theory is a proposition claiming to explain an aspect of the world, which is subject to testing and potentially falsification by hypothesis formation and experiment. Such a construction cannot be proven "true", because there is always the possibility that a counterexample may be found.

Music theory, on the other hand, is a collection of sets of rules which describe the culturally determined practice of people who create music in a particular culture during a particular period. Here, the notion of truth is slightly different, since music theorists readily acknowledge exceptions to their rules, and the arrangement is rather more like the statistical demonstrations found in the social sciences (though in a qualitative form). For example, tonal theory may be said to be one collection of rules, pitch class set theory might be said to be another. Music theory, more often than not, is motivated by didactic aims, to teach aspiring composers how previous composers have constructed their music. In a similar way, it gives performing musicians a practical means of communicating about structural aspects of a piece to be performed, and it can be used to convince sophisticated listeners of a certain interpretation of a piece.

Scientific theories of music, in between, try to explain, among other things, music as a cognitive activity and its relation to other cognitive faculties as well as to brain structures and functions, and to understand their relationships with other such structures and functions. Because music does not exist in the physical world, subtle methods from psychology must be used to treat it as rigorously as is possible.

And thus is the question of interest here posed: what is the relationship between music, music theory, scientific theories of music, and science?

A fundamental difference in kind between a music theory and a scientific theory is that a music theory does not *explain* things about music, but instead names and constrains them. Thus, for a very trivial example, a perfect cadence is a musical construct, which is *defined* in tonal music theory; however, having identified a perfect cadence in a piece of music, we are, in a sense, none the wiser — the theory describes, and permits structural analysis and naming, but it does little more: it allows music analysts to explain *what* in quasi-rigorous way (and the *what* may well be closely connected with the experience of the listener); it allows instances of music to be ruled out of certain classes (so Wagner's music is "demonstrably" not written by a Renaissance Italian) — but *why* is still an unanswered question, and must be left to informed introspection and speculation by music analysts and musicologists. Attempts to lead tonal music theory in a more systematic direction (*e.g.*, Schenker, 1925, *etc.*) have not borne particularly ripe scientific fruit, and attempts to match musical phenomena directly with rigorous abstract scientific methods, such as group-theoretic analysis (*e.g.*, Noll, 2010, this volume) leave us with a sense of elegant beauty and surprise, but no real *answers* — indeed, these analyses raise many more questions than they answer, when one looks at the wider view: why *should* an evolved cultural and psychological construct (such as we claim Music to be) exhibit such complex and mysterious regularities for extrinsic reasons? Unless this supplementary question is asked and ultimately answered, for example by elaborating Noll's closing speculation that musical mental activity may be related to implicit human expertise in mathematical dynamics, a match of this purely descriptive kind does not move understanding forward. Indeed, it risks the same idealist fallacy that arises in Chomskian linguistics, discussed below.

To say this is *not*, in any sense whatsoever, to denigrate the efforts of music scholars (including those cited above!) who have worked to make our understanding of music more precise and clear over the centuries. Wiggins (2007) argues, in the context of cognitive modelling, that a descriptive model is necessary before an explanatory one, and the same is true in this wider sense of theory: we need a precise, clear, reliable and complete description of a phenomenon (and a language in which to describe it) before we can realistically expect to explain it. Music theories (in their appropriate contexts) are exactly such a thing. They are "theoretical" in a sense that is complementary to "practical", and not in the scientific sense at all; they describe common practice in one or more aspects of music (*e.g.*, composition, performance, *etc.*), but they do not generally seek to explain why they are what they are.

The distinction between scientific theory and theory of music cast in this way begs a question to which the current paper is intended to propose an answer: if theory of music as it exists is not a scientific theory, then what *is* a scientific theory of music? If, as we argued above, music is not primarily in the acoustic or graphemic domains, but fundamentally an auditory, or psychological, thing, then we evidently need to study the related psychology. If, as we also argued above, it is actually something more than all of these domains put together, then we need to study not only each domain, but also the interactions between them.

In a weak sense, we might argue, conventional theories of music are themselves cognitive theories, because they describe something that only really has existence in a cognitive context. However, this idea is indeed weak in two ways: first of all, musical behaviour as observed in the world is only the tip of the iceberg of the psychological phenomenon, and therefore it is unlikely that one can achieve a truly explanatory theory on the basis of that, and, second, a cognitive model which aims only to describe does not, by definition, explain anything: any explanation has to be outside the theory. An example here might be the grouping rules of Lerdahl and Jackendoff (1983), which are derived by observation of musical structures in the area of music studied, combined with some background knowledge of Gestalt principles[2]. For example, one such rule, GPR2b states (in paraphase) that a relatively long duration is likely to end a group of notes. The rule is merely a statement of something that is clearly true of the musical domain in question, but it only allows us to say

(2) Indeed, the Gestalt principles themselves constitute a descriptive model, because the mechanisms behind them are not (yet) explained (Wiggins, 2007).

"there is likely to be a boundary here", and not to infer any reason for this other than "GPR2b says so". As an example on a larger structural scale, the *Grouper* system (Temperley, 2001) includes a parameter, set by hand, determining the length of a phase: 8 notes for Temperley's "Ottoman" collection, but 10 notes for the Essen folk database; the optimal value is determined by observation. Since minds have no facility for setting such a parameter, the setting, *ad hoc* in Temperley's model, must come from somewhere else — and Temperley's theory does not explain where or how, and so the model is not in fact explaining the process of segmentation. However, given a descriptive model such as the GPR rules or *Grouper*, we can make progress towards an explanatory model by beginning to study empirically how they can be applied, as do Bruderer *et al.* (2010, this volume). We will return to the fundamentally important question of grouping, in a later section.

Because music, and in particular musical structure, only has existence in the mind, the very notion of a scientific theory of Music, distinct from mind, is suspect[3]. That is to say, studying the extension into the world of the phenomenon — the sounds and the notes on paper — is only to study *its effect on the world*, and not the thing itself. To study the thing itself, we need access to the implicit, or tacit, knowledge used by music analysts — the structures that are *inferred* and *experienced* by listeners and other active musicians — and to the processes that build them. That knowledge is only available in rigorous form to psychological study, and so we propose that, if we are really to understand this thing which is *so* important to human culture, we need to understand not only Music, *qua* "cultural construct", but also *Musical Behaviour*, interpreting "behaviour" to mean all levels of processing and generation, cognitive and embodied, from auditory harmonic fusion to folk dance to the composition of symphonies. Otherwise, one studies an incomplete phenomenon, and success is unlikely.

Conventional music theory succeeds precisely *because* it makes no attempt to go beyond description (and it is nevertheless useful, because it describes and communicates things musicians want to know about), but it is doomed to a future of mere description, as opposed to more interesting *prediction* because of its purely descriptive nature. Tsougras (2010, this volume) shows how rules from GTTM, developed for mainstream Western art music before 1900, can be applied to certain pieces of 20th century music in a non-mainstream style. He performs metrical, grouping, reductional, and prolongational analyses on the *44 Greek miniatures for piano* of Yannis Constantinidis, a collection of contemporary modal pieces. The analysis starts from the assumption that the cognitive-analytic principles of GTTM are universal, but it becomes clear that some preference rules need to be re-written and complemented with additional rules to account for the stylistic differences and

(3) Again, we must emphasise that this is *not* to suggest that non-scientific theories of music are problematic by nature.

the context of modal music. So the work shows that the analytic rules proposed in GTTM can to some degree be applied for music analysis outside the domain of the original theory. On the other hand, it constitutes a traditional music analysis where the epistemological focus is limited to the pieces of music analysed, and thus cannot be used to argue for the validity, cognitive or otherwise, of GTTM rules.

What, then, is GTTM? Lerdahl and Jackendoff's own presentation draws on perceptual and cognitive principles, but does not make very strong claims about its cognitive nature: the chapter devoted to this issue is speculative, and makes a more modest (and clearly defensible) concluding claim:

> Thus music theory is by no means a curious side branch of cognitive science. We believe we have shown that it can provide central evidence toward a more organic theory of mind.
>
> (Lerdahl & Jackendoff, 1983, p. 332)

Nevertheless, GTTM has initiated and heavily influenced important interdisciplinary music research in the last 25 years. Among the domains where it has been influential are formal music theory, grammatical models of music cognition, and rule-guided music analysis (of western art music). To understand the status of GTTM better, it is necessary to go back to its roots.

RE-PARSING[4] CHOMSKY

IDEALIST PHILOSOPHY IS NOT ALWAYS IDEAL

Lerdahl and Jackendoff (1983, p. ix) trace the roots of the GTTM back to the famous MIT seminar on music, linguistics, and aesthetics, which in turn was inspired by Bernstein's 1973 Harvard lectures (Bernstein, 1976) which, among other things, "advocated a search for a 'musical grammar' that would explicate human musical capacity." We note here that this implicitly places the authors in opposition to our claims above: the implicit assumption is that a grammar capturing the surface form of music (as visible in the world), perhaps associated with annotations somehow equivalent to the parts of speech in linguistics in some form of tree is adequate to "explicate the human musical capacity". This is self-evidently incomplete: implicit in the application of a grammar to any data is a *process* though which the data is *parsed according to the grammar* — the grammar (or any other set of rules) on its own will *do* nothing, and therefore can produce no output. Attempts to implement GTTM (*e.g.*, Curry & Wiggins, 1999; Hamanaka *et al.*, 2007) have foundered on this rock, being forced to add decision processes: the mechanism is underspecified, and relies on expert knowledge by a human user.

(4) **parse:** "examine or analyze minutely" (meaning 3, New Oxford American Dictionary).

In defence of Lerdahl and Jackendoff, in historical context, it is only fair to note that Chomsky himself made the same omission, deliberately and explicitly:

> Linguistic theory is concerned primarily with an ideal speaker-listener, in a completely homogeneous speech-communication, who knows its (the speech community's) language perfectly and is unaffected by such grammatically irrelevant conditions as memory limitations, distractions, shifts of attention and interest, and errors (random or characteristic) in applying his knowledge of this language in actual performance.
>
> (Chomsky, 1965, p. 3)

Thus, Chomsky dismisses cognitive mechanism as "irrelevant", while implicitly assuming, but not defining, an ideal mechanism by which a grammar can be used to parse his "ideal" language. These notions have, of course, been undermined over time by the development of more modern linguistic methods, such as quantitative and corpus linguistics and natural speech processing (*e.g.*, Manning & Schütze, 1999). Central concepts to the Chomskian view of language, such as grammaticality, the role of change and development of a language, and the emphasis of linguistic performance *vs.* competence have been discussed, refined, and in some cases changed radically as part of of the development of modern, evidence-based linguistic approaches (*e.g.*, McEnery & Wilson, 2001).

Chomsky's way of thinking, though, chimes strongly with the traditional music-theoretic approach: observing an idealised version of the phenomenon, and treating it as though it had its own existence, independent of the entities exhibiting the behaviour that produced it — Chomsky's ideal speaker-listener certainly does not exist.

One is reminded of high-school physics examinations, beginning "A perfectly round, frictionless ball is resting on a frictionless table, tilted at an angle of 45°…" — and the idealisations may be made for the same reasons of tractability. However, in high-school physics, these are *approximations*; they do not fundamentally change the nature of the phenomenon, but merely simplify it slightly. In music and language, omitting the process from the account is akin to missing out gravity in the physics question. Without it, nothing happens.

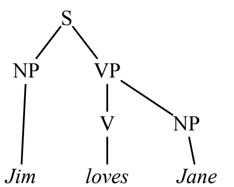
Grammar and parsing

Chomsky's production-rule-based approach to the description of language is by no means the only one, though it has (detrimentally) eclipsed most others. For example, Woods (1970) used *augmented transition networks* to parse English language; and Steedman (1999) has very successfully developed and promoted another formalism, *categorial grammar*, founded on the extended λ-calculus of Lambek (1958). Both of these approaches have explicit mechanism: the ATN is a network (enhanced with a memory store) which must be *traversed*, choices being made, where necessary, on the basis of input symbols in sequence; categorial grammar has its notion of function application (or, in Lambek terms, β-reduction). Both of these formalisms work, in a linguistic sense, bottom-up — but note that this is a different usage of this term (and the complementary "top-down") from that prevalent in psychology: it means that the grammar is to some degree *lexicalised* (Steedman, 1999), in that structures (and meaning) are composed by grouping words together by virtue of themselves, as opposed to starting out by "producing" a sentence, and then deciding to "produce" whatever syntactic category a sentence is defined to start with, and so on, which is the top-down way.

Chomsky's statement of grammar is in fact agnostic as to parsing mechanism (recall that it is intended to describe idealised language), but the statement of the actual grammars strongly suggests two things: first, top-down processing, starting with some high-level goal (usually, a sentence, S); and second, a recursive process (because rules can reuse themselves), which leaves a trace through the grammar in the shape of a *parse tree*. All this is illustrated by Figure 2, in comparison with a categorial grammar.

(a) Generative Formulation

$$\begin{array}{rcl} \mathrm{S} & \rightarrow & \mathrm{NP}\ \ \mathrm{VP} \\ \mathrm{VP} & \rightarrow & \mathrm{V}\ \ \mathrm{NP} \\ \mathrm{NP} & \rightarrow & \{\textit{Jim}, \textit{Jane}, \ldots\} \\ \mathrm{V} & \rightarrow & \{\textit{loves}, \ldots\} \end{array}$$



(b) Categorial Formulation

$$\begin{array}{rcl} \textit{Jim} & := & \mathrm{np} \\ \textit{Jane} & := & \mathrm{np} \\ \textit{loves} & := & (\mathrm{s}\backslash\mathrm{np})/\mathrm{np} \\ \ldots & & \\ \mathrm{X/Y + Y} & \rightarrow & \mathrm{X} \\ \mathrm{Y + X\backslash Y} & \rightarrow & \mathrm{X} \end{array}$$

$$\begin{array}{c} \textit{Jim} + \textit{loves} + \textit{Jane} \\ \Downarrow \\ \mathrm{np} + (\mathrm{s}\backslash\mathrm{np})/\mathrm{np} + \mathrm{np} \\ \Downarrow \\ \mathrm{np} + \mathrm{s}\backslash\mathrm{np} \\ \Downarrow \\ \mathrm{s} \end{array}$$

Figure 2.

(a) A very simple Chomskian grammar and the parse tree for "Jim loves Jane", as compared with (b) a categorial equivalent, including the combination rules, and its derivation. Note the clear structural relationships between the two, but also that (a) has the implication of parsing from top to bottom, while (b) is explicitly parses up, from the words, and discovers *a sentence. Abbreviations: n is noun, v is verb, p is phrase; X and Y are variables.*

Tsoulas (2010, this volume) addresses this issue by showing how relatively modern methods of bottom-up lexical grouping, based on one simple "merge" operation, can do without the kinds of syntax trees that Chomsky espouses, and relating this to music. Wiggins (1991) shows how to do this in a categorial context, using an extra categorial combination rule, *protraction*, which allows strictly incremental processing of syntax, like the merge rule. However, the result of doing so is that the grouping encoded in the trees is lost, and needs to be recorded in some other way.

Notwithstanding the fact that the implicitly top-down, idealised processing of Chomsky and the explicitly bottom-up categorial approach exhibit a clear correspondence, the former tends to lead, in the literature, to the assumption of tree structures while the latter (especially the incremental versions) explicitly exclude trees as an intermediate or final cognitive representation: as can be seen in the current volume, the effect of incremental processing is to force the tree to degenerate into a sequence, which means that it can no longer account through its own structure for unbounded dependency. This raises a question of how grouping structure is determined cognitively, and we discuss this in the next section.

Competence, performance and recursion

Chomsky introduces a distinction between *competence* — the language understandable by the ideal listener — and *performance* — the language generable by the speaker; presumably this relationship is somewhat akin to the two classes of grammar that Baroni (2010, this volume) describes. Again, here, it is perhaps surprising that Chomsky excludes "memory limitations" as an "irrelevant" point. In this section, we aim to show that they are very relevant indeed.

A fundamentally important claim of Chomskian linguistics is that human linguistic grammars are *recursive*:

> To be precise, we suggest that a significant piece of the linguistic machinery entails recursive operations…
>
> (Fitch *et al.*, 2005, p. 182)

This word is borrowed from mathematics, where it has a very precise meaning: a process or function which is defined in terms of itself. Recursion is an extremely powerful operation, so much so that it allows the generation of infinitely many and infinitely long strings of language from finite, even tiny, grammars. Figure 3 shows how this is possible.

The *infinite* quality of recursive processes is important in mathematics, when it is sometimes necessary to be able to conceptualise infinite objects in finite terms. However, in linguistics, we can be absolutely certain that no human utterance will be infinitely long, for self-evident practical reasons; it is meaningless, therefore, to suggest that an idealised one may be so. What is important, in fact, is that a grammar

(a) $X \rightarrow g$ (b) $X \rightarrow g\,X$ (c) $\begin{array}{l} X \rightarrow g \\ X \rightarrow g\,X \end{array}$

Figure 3.

Three very small grammars, two of which generate infinite structure: (a) produces a set containing exactly one finite string, {g}; (b) produces a set containing exactly one infinite string, {…ggg…}; (c) produces a set containing an infinite number of finite strings and one infinite one {g,gg,ggg,…,…ggg…}. The presence of a single recursive rule (X defined in terms of X) makes the difference.

may generate sentences of *arbitrary*, *finite* length, so that linguistic expression is not restricted; Chomsky is concerned that his grammars will generate all the sentences that *might* be generated, regardless of whether they have been or will be. It is philosophically unclear whether a language with a finite lexicon existing in a finite world needs to be able to generate infinitely many sentences, so opting for the more generous formalism with respect to this aspect is reasonable. Whatever, the claim of infinitude (Pullum & Scholz, 2010) is not a claim that can be made about empirical (*i.e.*, real) linguistics and its associated mechanisms, because the entities engaging in it are finite. Therefore, there is no reason to suppose that the philosophical idealisation bears any relation to reality whatsoever.

What is particularly important about so-called recursion in grammars is the facility to embed phrases, one inside another. For example, the sentence

The man who bit the dog was barking.

contains two basic noun phrases ("the man" and "the dog"), but the latter is *subordinated* to the former, by virtue of the relative pronoun, "who", producing a third composite noun phrase, "the man who bit the dog". A naïve grammar capable of parsing this and the corresponding tree are shown in Figure 4.

This capacity to embed is clearly very important to human expression, and it generates important features in many languages. It is a desirable feature of languages described by these grammars that embedding is strictly nested (like GTTM's GPRs), with embedded phrases and clauses neatly contained within their superordinate phrase and lined up in the "correct" order, as, for example, with the very strict verb-final subordinate clauses of German; unfortunately, this is not the case in various languages (*e.g.*, Dutch, Japanese: Steedman, 2000). This is just one of many examples where attempts to make general claims about language syntax end in collections of more or less *ad hoc* rules.

Much more importantly, in the same way that human processing and generation are biologically limited to finite utterances, so human memory is limited to quite a small number of concurrent embeddings, in the case where information has to be

$$
\begin{aligned}
\text{S} &\rightarrow \text{NP VP}\\
\text{VP} &\rightarrow \text{Copula Adj}\\
\text{VP} &\rightarrow \text{V NP}\\
\text{NP} &\rightarrow \text{Det N}\\
\text{NP} &\rightarrow \text{Det N RelC}\\
\text{RelC} &\rightarrow \text{RelP VP}\\
\text{Det} &\rightarrow \{\textit{the}, \ldots\}\\
\text{N} &\rightarrow \{\textit{man, dog}, \ldots\}\\
\text{Copula} &\rightarrow \{\textit{was}, \ldots\}\\
\text{V} &\rightarrow \{\textit{bit}, \ldots\}\\
\text{RelP} &\rightarrow \{\textit{who}, \ldots\}\\
\text{Adj} &\rightarrow \{\textit{barking}, \ldots\}
\end{aligned}
$$

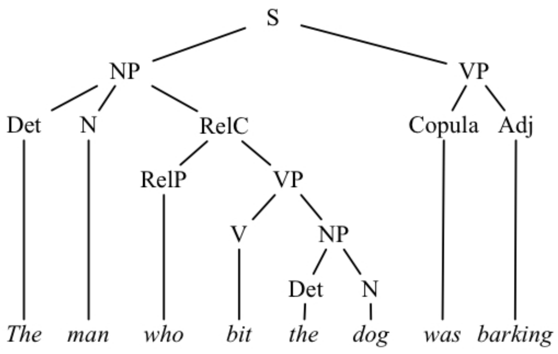


Figure 4.

A grammatical example, showing the embedding of one noun phrase (NP) inside another, by relative subordination. Note that this grammar is not intended to be representative of relative subordination in general, but is designed for the purposes of example.

carried forward during an unbounded dependency. To see this, consider the sentences in Table 1.

Table 1

A sequence of sentences which are progressively harder to parse implicitly, because of their progressively heavier centre-embedded relative clauses

1.	"The man I saw yesterday was painting the fence."
2.	"The man I saw who was painting the fence yesterday dropped his brush."
3.	"The man I saw who was painting the fence in the park yesterday dropped his brush."
4.	"The man I saw who was painting the fence in the park by the river yesterday dropped his brush."
5.	"The man I saw who my friend said was painting the fence in the park by the river yesterday dropped his brush."
6.	"The man I saw who the friend that I have known for many years said was painting the fence in the park by the river yesterday dropped his brush."
7.	"The man I saw who the friend with whom I was at school and whom I have therefore known for many years said was painting the fence in the park by the river yesterday dropped his brush."

By the time we get to sentence 7, it becomes nearly impossible to remember who drops the brush: is it my friend or the man of whom he speaks? This is a simple demonstration of a principle well-established in empirical linguistics: embedding *in practice* is limited to 3 or 4 levels in writing and fewer in speech (*e.g.*, Karlsson, 2007). There is no infinite data structure to do the house-keeping — indeed, the available data structure is pitifully small, by comparison with the stack in the corresponding computational parser (which, of course, is still finite).

One should not deny the possibility of recursion in cognition on the basis of this argument alone (Fitch *et al.*, 2005, claim that recursion is a generic cognitive ability setting humans apart from other species). The point here is that infinite recursion is not necessary for human language, and therefore language cannot be used as support for recursion as a generic cognitive ability. The consequence of this is that the Chomsky (1965) hierarchy of grammar types is irrelevant to human language[5], because centre-embedding types of grammar (Context-Free and above) can be approximated by Finite State grammars so long as infinite recursion is not allowed (Mohri & Nederhof, 2000).

Syntax, semantics and the meaning of "meaning"

Chomsky begins his argument about language with formalisations of sentence structure, rather than formalisations of meaning (indeed, we would argue, this is a another symptom of viewing the language as an artefact in its own right, as opposed to a communication medium). While there can be no denying that preferred syntactic forms exist, they are not strictly necessary for effective, basic linguistic communication, but rather facilitate and sometimes disambiguate it, as any beginning speaker of a non-native language knows: "Me Tarzan. You Jane." is every bit as comprehensible in context as the more typically Victorian, "Dr Livingston, I presume?"; and the cutely ill-formed syntax of Yoda the Jedi Master also comprehensible is[6]. This is because the semantic content is primary, and not the syntax. Syntax becomes more important where deixis is not possible, or where abstract ideas are involved, but as the international scientific community shows, perfect English syntax is not necessary to communicate really quite complicated ideas. This strongly counter-Chomskian point is important because, if an analogy is to be drawn between GTTM and Chomskian theory, we need to understand which parts go where.

It is sometimes assumed that, because both syntax and semantics exist in language, then the same must be true in music. Wiggins (1998) argues against this assumption, and suggests that, while it is not unreasonable to talk about "meaning" in music, there is little, if any, of what linguists would call "semantics" in most music, and that in musical studies, in fact *syntax* is primary. Lerdahl and Jackendoff implicitly concur with this standpoint: they make no attempt to explicate any

(5) Fortunately, it is nevertheless very relevant to computer science.

(6) If sense this does not make, *Star Wars* see you must.

semantics in their data and the word does not even appear in their index: its usage in the book is strictly limited to linguistic semantics. Steedman (1996), on the other hand, proposes that the semantics of jazz chords is to be found in their harmonic implication. If this is so, then it is a different kind of semantics from the kind that includes reference to real world entities and relationships between them, though one can clearly see the *analogy* between these and harmonic relationships; perhaps it is more like the low-level cognitive, experiential semantics of Gärdenfors (2000) which can be applied to music (Forth *et al.*, 2008). What is clear, however, is that musical semantics, if it exists, does not normally refer to the physical world in the same way as language, and on the few occasions when it does so, it does so metaphorically or (in the limit) onomatopoeically. Nevertheless, music has its own notion of reference, which is very important; we return to this below.

When we import linguistic theory into music, though, there is still another consideration: the relationship between syntax and semantics in Chomskian linguistics (Chomsky, 1965) that seems to have been misleading in the past. The confusion is that Chomskian "syntactic" categorisation (such as "noun phrase" and "adjective") is just as *semantically* motivated as the distinction between the individual meanings of, for example, nouns (such as "elephant" and "chair"); the difference is not in nature, but in level of abstraction. This becomes evident when one considers linguistic sub-categorisation, for example, of nouns into singular and plural: there are, of course, syntactic markers which *correspond* with each sub-categorisation, but each so-called "syntactic category" is actually (in correspondence with) a *semantic* one; and the semantic categories form a semi-lattice, rather than a simple tree, when one places them in partial order of containment. What are generally called the "syntactic categories" — the high-level ones, such as "noun phrase" and "verb phrase" — are in reality *semantic* categories which are so abstract that removal of their last defining feature leaves them mutually indistinguishable; syntax resides not in the categories themselves, but in the (learned, not absolute) rules defining the sequences in which the categories are expected to appear, and in the morphology which identifies words with their intended (sub-)categories. The high-level semantic (sic) categories are inter-linguistically stable because they are firmly grounded in the world (things *vs.* actions); however, their syntactic equivalents are not: they are driven by the need to make the semantics clear.

The order of the main clause in languages is important: there is a subject (S) which refers to the agent of the clause, a verb (V), which expresses the relation between the agent and the patient, and then an object (O) which refers to the patient — note that the thing common in all languages is the semantics: the agent, patient and relation. However, languages exist which exhibit *each* possible syntactic order produced by permuting these three structures[7] (SVO, SOV, *etc.*), though some are

(7) Examples are: SVO, English; SOV and OSV, Japanese; OVS, Tamil; VSO, Welsh; VOS, Fijian.

more common than others. In terms of Ockham's razor, a better theory would only require the rule that a transitive verb needs both a subject and an object; the morphology of the words would mark these semantic subcategories; and we simply require a small amount of memory to make sure that the two sememes produced by the subject and object are stored until the verb arrives in SOV languages, or the converse in VSO language, such as Welsh. But this kind of simpler explanation is impossible in a theory that limits itself to the artefact at the expense of the process.

The upshot of our argument is that the tree structures generated by generative language theories are more about semantic grouping than they are about syntactic grouping: the syntax *serves* the semantics, but has no function beyond this. Therefore, since we have argued that music is without the main kind of linguistic semantics, this kind of tree might be said to have even less to do with music than it has with language. Evidence for this resides in the fact that the basic structural units of language (whether one calls them syntactic or semantic) have no equivalents in music, and need to be dropped when one tries to import linguistic methods to music (Bod, 2001).

Baroni (2010, this volume) pursues Lerdahl's (1988) notion of different but related musical grammars in a composer and listener, both engaging with the same piece of music. Because this distinction has not been elaborated by Lerdahl himself, Baroni makes it very obvious by considering 20th century avant-garde music. Here, compositional grammars explicitly serve composers to create works; composers often generate and apply new grammars for each new piece. Listening grammars are often explicitly separated from these compositional grammars and serve listeners to make sense of a new work. Baroni also stresses that, while grammatical systems may overlap, music grammars are always style-specific — a fact that is implicit in GTTM and its follow-up theories. Thus, in Baroni's conception, musical grammars are always grammars of musical style[8], and there is a fundamental difference between musical grammars and linguistic grammars: Chomskian "deep structure" does not exist in music. This, Baroni says, explains why music is not translatable: a musical structure in one style and context does not have an exact correspondence in a different musical style. We interpret this as supporting our argument that music has no (linguistic-style) semantics and Steedman's suggestion that musical semantics exists at an absolute (untranslatable) perceptual level.

A very different approach to musical meaning, opening a wider perspective on understanding music and human musical interaction is suggested by Leman (2007, 2010, this volume). Leman emphasises the importance of corporeal and collaborative semantics where meaning is introduced in the music perception/cognition process

(8) In our terms, this would mean that theory is strictly descriptive, and that to make it explanatory, one would have to say where different stylistic grammars come from and give appropriate process models.

by the alignment of the human body with musical actions and other individuals engaging in the same musical activities, in ways analogous with theories of symbol-grounding in language (*e.g.*, Bleys *et al.*, 2009). Leman acknowledges that a lot of evidence has been accumulated, from experimental psychology, computational approaches and neuroscience, in the context of structural semantics and following GTTM as probably the most influential paradigm in music cognition research since the early 1980s. This evidence demonstrates structural regularities and invokes notions of musical grammar — but consistently has difficulties when musical syntax is "wrong" (*i.e.*, not compliant with the specific grammar at hand) or very different from Western music. We concur with Leman that, although we accept the potential value of quasi-syntactic analysis of music in the purely music-analytical context, there are severe limitations as to what it can tell us at the cognitive level, and therefore we must question to what extent the approach really explicates music. Indeed, as its authors agree, GTTM is in fact GTWTM(1750-1900): the Generative Theory of Western Tonal Music from 1750-1900, and this leads us to another point about the *mutable* nature of musical and linguistic theories, which we address in the next section.

Ultimately, the arguments above, and others in the literature, have strongly undermined the top-down, structuralist view of language that was prevalent in the late 1970s and which influenced the approach taken in GTTM. Where does this leave us, in relation to GTTM's status as a cognitive theory of music?

Breaking the rules

Mutability of language and music

Any theoretical system attempting to model a phenomenon that is subject to change across time, such as language or music, needs to account for that mutability in a principled way. Language changes spontaneously as society develops; and a statement of what a particular style of music "is" serves as an immediate imperative challenge to any self-respecting composer who identifies with that style. There is plenty of high-quality research on learning systems which are capable of inferring grammatical rules, though no artificial system currently exists which can really efficiently learn grammars more powerful than Context-Free (Chomsky Type 2) grammars. At a philosophical level, this should lead us to question the notion of Universal Grammar in human language, which needs to learn grammars more powerful than Type 2, if one believes in recursion as a primary feature of language: the evident (and deeply studied) mutability of language and music inevitably raises insistent questions about the cognitive nature of the systems underlying them, because viewing any example as an absolute, and seeking to generalise from it without considering the source of the phenomenon is to deduce from mostly incomplete premises. Like Linguistics, Music Theory, as a whole, has avoided this problem by focussing on more or less fixed, isolated and separate bodies of data, albeit large ones — GTTM is a prime example of this, and Baroni's and Leman's remarks emphasise our point. When one seeks generality, however, problems arise.

What this means for modelling approaches is that one needs to think not of a model of the observed behaviour or artefact, per se, but of a model of how that behaviour or artefact is generated and can change. This is then an explanatory model, in the terms of Wiggins (2007), because it must, by nature, be able to predict future changes (and can therefore be tested) as well as helping to account for how the observed phenomenon arises in the first place. Ideally, such a model will appeal to processes that are known to be available in its context (*viz.*, known cognitive processes, in the current one).

Statistical modelling of musical behaviour: A *parallel* with GTTM?

One such model is our own IDyOM model (Pearce & Wiggins, 2006; Potter *et al.*, 2007; Pearce *et al.*, 2008; Wiggins *et al.*, 2009; Pearce *et al.*, 2010). This began as a model of style acquisition and, in particular, melodic pitch prediction (Pearce, 2005), but has since proven to serve other related purposes. To make the distinctions between the various aspects of its behaviour clear, we have coined the term *meta-model* to name a computational model which is able to predict behaviour in a domain related to, but not the same as, that for which it was designed (Wiggins *et al.*, 2009), in addition to its original task. Thus, we add support for the original model as being at some level veridical (Honing, 2006). In this case, the very same model is used to predict segmentation of melodies in various styles (Potter *et al.*, 2007; Pearce *et al.*, 2008).

This process of segmentation, is, of course, equivalent to grouping, and this brings us back full circle to the fundamental motivation of GTTM: the explication of musical structure in terms of grouping at various levels, which then motivates and enables analysis in terms of time-span reduction and parallelism (in fact, Lerdahl and Jackendoff choose not to explore this last aspect in their original presentation). Following Steedman's idea that harmony might be semantics, this gives us the closest analogy with the Chomskian world yet: grouping can be seen as syntax; reduction can be seen as explication of harmonic "meaning".

The IDyOM research does not claim to explicate a particular style of music, but is focused on a fundamental property of music which seems to be fundamental to all musical cultures: the perception of sequences of sonic objects in time. It would be repetitious to explain the detail again here: suffice it to say that the system performs acceptably at its various tasks, and descriptions and evaluation of its behaviour are published in the papers cited above. However, there is an important symmetry with GTTM that is perhaps surprising: GTTM uses grouping to underlie parallelism (at least in principle); whereas IDyOM uses parallelism to underlie grouping.

Lartillot (2010, this volume) gives an elegant exposition of his own model of parallelism, extending GTTM as its progenitors suggested; IDyOM, too, implicitly includes, not a *model* of parallelism, but a process based on the fact that parallelism is a *sine qua non* of memory for sonic sequences: retrieval of such memories can only be based on a process which has the ultimate effect of matching along the paired sequences (even if the corresponding neural process is not exactly this). The model works by *compressing* the memory of sequences heard (in much the same way as the zip program found on most modern computers compresses sequences of letters), based on the hypothesis that brains compress information from the world in order to keep it manageable. To do so, it finds common subsequences in the data to which it is exposed, allowing things that appear many times to be represented just once. It then uses the memory of what came after previously-experienced subsequences to predict what comes next, note by note, during the experience of a new sequence. We suggest that this process of finding related substructures in memory is the closest musical phenomenon to anaphoric reference in language; in any case, the recognition of such substructures, large and small, is cited by many authors as an important organising principle in music: Lerdahl and Jackendoff's parallelism.

Like GTTM, IDyOM is inspired by similar models from linguistics (Saffran *et al*., 1996, 1999). These ideas are enhanced by the multiple-viewpoint data representation of Conklin and Witten (1995), which allows the principled combination of different models of the various parallel perceptual parameters in music. Unlike GTTM, IDyOM is inspired by the underlying process enabling the phenomenon (learning and perception) rather than the phenomenon itself (music). We have recently demonstrated its generality by taking it back to linguistics, where it performs surprisingly well, with no special adaptations to the different kinds of sequence found there (Wiggins, 2010).

Conclusion

We suggested at the start of this paper that music does not exist, in and of itself, and we have given an argument as to why: it cannot exist unless a mind is implicated. We have also argued that neither music nor language can be studied as pure surface forms, because the cognition of both produces information which is not contained in the surface form. We have suggested that this has consequences for methodology in this kind of study, and that the Chomskian approach to linguistics is doomed to failure on account of its overly positivistic approach to the linguistic phenomenon itself, though this does not discount an appropriately positivistic approach to linguistic or musical *processes*.

What we have not yet justified is our initial claim that "Music Theory is a figment of the imagination". This is not, as it might seem, a purely dialectic polemic; rather, it is a positive statement about the nature of music theory. Music Theory is the very best kind of folk psychology: it is principled, carefully worked out, and based in what is as close as can be to agreed "truth". Music Theory is not a figment of the imagination in the dismissive sense of that phrase, but it aims to, and to a large extent does, capture and explicate the figment of the human imagination that constitutes the capacity to hear, listen to and understand the magnificent and uniquely human phenomenon of Music.

And thus we see the value of extending Music Theory explicitly into cognition: for the figment, in truth, is the very nature of the Music.

Acknowledgements

We are grateful to colleagues in the Intelligent Sound and Music Systems group at Goldsmiths for conversations which have contributed to much of the thinking presented here, and to Irène Deliège and Moreno Andreatta for inviting us to contribute to the workshop at IRCAM in January 2008, which gave rise to this volume. Above all, we acknowledge the spectacular contribution of Fred Lerdahl and Ray Jackendoff to music theory and music cognition studies with the Generative Theory of Tonal Music.

Address for correspondence:
Geraint A. Wiggins
Goldsmiths, University of London
Department of Computing
25 St James
New Cross
London SE14 6NW UK
e-mail: g.wiggins@gold.ac.uk

• References

Aucouturier, J.-J. & Pachet, F. (2004). Improving timbre similarity: How high's the sky? *Journal of Negative Results in Speech and Audio Sciences*, *1*(1).

Babbitt, M. (1965). The use of computers in musicological research. *Perspectives of New Music*, *3*(2), 74-83.

Baroni, M. (2010). GTTM and post-tonal music. *Musicæ Scientiæ*, Discussion Forum 5 "Lerdahl and Jackendoff's GTTM: 25 years on", pages 000-000.

Bernstein, L. (1976). *The Unanswered Question*. Cambridge, MA: MIT Press. Based on the 1973 Harvard Lectures.

Bleys, J., Loetzsch, M., Spranger, M., & Steels, L. (2009). The grounded color naming game. In *Proceedings of 18th IEEE International Symposium on Robot and Human Interactive Communication, Ro-man*.

Bod, R. (2001). Memory-based models of melodic analysis: Challenging the gestalt principles. *Journal of New Music Research*, *30*(1), 27-37.

Bruderer, M. J., McKinney, M. F., & Kohlrausch, A. (2010). The perception of structural boundaries in polyphonic representations of Western popular music. *Musicæ Scientiæ*, Discussion Forum 5 "Lerdahl and Jackendoff's GTTM: 25 years on", pages 000-000.

Chomsky, N. (1965). *Aspects of the theory of syntax*. Cambridge, MA: MIT Press.

Conklin, D. & Witten, I. H. (1995). Multiple viewpoint systems for music prediction. *Journal of New Music Research*, *24*, 51-73.

Curry, B. & Wiggins, G. A. (1999). A new approach to cooperative performance: A preliminary experiment. *International Journal of Computing Anticipatory Systems*, *4*.

Fitch, W. T., Hauser, M. D., & Chomsky, N. (2005). The evolution of the language faculty: Clarifications and implications. *Cognition*, *97*, 179-210.

Forth, J., McLean, A., & Wiggins, G. A. (2008). Musical creativity on the conceptual level. In *Proceedings of the International Joint Workshop on Computational Creativity*, Madrid, Spain.

Gärdenfors, P. (2000). *Conceptual Spaces: the geometry of thought*. Cambridge, MA: MIT Press.

Hamanaka, M., Hirata, K., & Tojo, S. (2007). Implementing "a generative theory of tonal music". *Journal of New Music Research*, *35*(4), 249-77.

Honing, H. (2006). Computational modeling of music cognition: a case study on model selection. *Music Perception*, *23*(5), 365-76.

Karlsson, F. (2007). *Journal of Linguistics*, *43*(2), 365-92.

Kraemer, D. J., Macrae, N. C., Green, A. E., & Kelley, W. M. (2005). Musical imagery: Sound of silence activates auditory cortex. *Nature*, *434*, 158.

Lakatos, I. (1978). *The Methodology of Scientific Research Programmes: Philosophical Papers*. Cambridge, UK: Cambridge University Press. (Edited by John Worrall and Gregory Currie).

Lambek, J. (1958). The mathematics of sentence structure. *American Mathematical Monthly*, *65*, 154-70.

Lartillot, O. (2010). Towards a generative theory of musical parallelism. *Musicæ Scientiæ*, Discussion Forum 5 "Lerdahl and Jackendoff's GTTM: 25 years on", pages 000-000.

Leman, M. (2007). *Embodied Music Cognition*. Cambridge, MA: MIT Press.

Leman, M. (2010). An embodied approach to music semantics. *Musicæ Scientiæ*, Discussion Forum 5 "Lerdahl and Jackendoff's GTTM: 25 years on", pages 000-000.

Lerdahl, F. (1988). Tonal pitch space. *Music Perception*, *5*(3), 315-50.

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

Manning, C. D. & Schütze, H. (1999). *Foundations of Statistical Natural Language Processing*. Cambridge, MA: MIT Press.

McEnery, T. M. & Wilson, A. (2001). *Corpus Linguistics*. Edinburgh: Edinburgh University Press, 2nd edition.

Mohri, M. & Nederhof, M.-J. (2000). Regular approximation of context-free grammars through transformation. In J.-C. Junqua & G. van Noord (eds), *Robustness in Language and Speech Technology*, pp. 251-61. Kluwer Academic Publishers.

Noll, T. (2010). Two notions of well-formedness in the organization of musical pitch. *Musicæ Scientiæ*, Discussion Forum 5 "Lerdahl and Jackendoff's GTTM: 25 years on", pages 000-000.

Pearce, M. T. (2005). *The Construction and Evaluation of Statistical Models of Melodic Structure in Music Perception and Composition*. PhD thesis, Department of Computing, City University, London, UK.

Pearce, M. T., Herrojo Ruiz, M., Kapasi, S., Wiggins, G. A., & Bhattacharya, J. (2010). Unsupervised statistical learning underpins computational, behavioural and neural manifestations of musical expectation. *NeuroImage*, *50*(1), 303-14.

Pearce, M. T., Müllensiefen, D., & Wiggins, G. A. (2008). Melodic segmentation: A new method and a framework for model comparison. In *Proceedings of ISMIR 2008,* pp. 89-94.

Pearce, M. T. & Wiggins, G. A. (2006). Expectation in melody: The influence of context and learning. *Music Perception*, *23*(5), 377-405.

Popper, K. (1959). *The Logic of Scientific Discovery*. Abingdon, UK: Routledge. Translation of German original, 1934.

Potter, K., Wiggins, G. A., & Pearce, M. T. (2007). Towards greater objectivity in music theory: Information-dynamic analysis of minimalist music. *Musicæ Scientiæ*, *11*(2), 295-324.

Pullum, G. K. & Scholz, B. C. (2010). Recursion and the infinitude claim. In H. van der Hulst (ed), *Recursion in Human Language*. Berlin: Mouton de Gruyter. In press.

Saffran, J. R., Aslin, R. N., & Newport, E. L. (1996). Statistical learning by 8-month-old infants. *Science*, *274*, 1926-28.

Saffran, J. R., Johnson, E. K., Aslin, R. N., & Newport, E. L. (1999). Statistical learning of tone sequences by human infants and adults. *Cognition*, *70*, 27-52.

Schenker, H. (1925). *Das Meisterwerk in der Musik, Volume I*. Munich: Drei Maksen Verlag.

Steedman, M. J. (1996). The blues and the abstract truth: Music and mental models. In *Mental Models In Cognitive Science*, pp. 305-18. Mahwah, NJ: Erlbaum.

Steedman, M. J. (1999). Categorial grammar. In *The MIT Encyclopedia of Cognitive Sciences*. Cambridge, MA: MIT Press.

Steedman, M. J. (2000). *The Syntactic Process*. Language, speech, and communication. Cambridge, MA: MIT Press.

Temperley, D. (2001). *The Cognition of Basic Musical Structures*. Cambridge, MA: MIT Press.

Tsougras, C. (2010). The application of GTTM on 20th century modal music: Research based on the analysis of Yannis Constantinidis's "44 Greek miniatures for piano". *Musicæ Scientiæ*, Discussion Forum 5 "Lerdahl and Jackendoff's GTTM: 25 years on", pp. 000-000.

Tsoulas, G. (2010). Computations and interfaces: Some notes on the relation between the language and the music faculties. *Musicæ Scientiæ*, Discussion Forum 5 "Lerdahl and Jackendoff's GTTM: 25 years on", pp. 000-000.

Wiggins, G. A. (1991). *An Adaptable Formalism for the Computational Analysis of English Noun Phrase Reference*. PhD thesis, Department of Artificial Intelligence, University of Edinburgh.

Wiggins, G. A. (1998). Music, syntax, and the meaning of "meaning". In *Proceedings of the First Symposium on Music and Computers*, Corfu, Greece.

Wiggins, G. A. (2007). Models of musical similarity. *Musicæ Scientiæ,* Discussion Forum 4a, 315-37.

Wiggins, G. A. (2009). Semantic Gap?? Schemantic Schmap!! Methodological considerations in the scientific study of music. In *Proceedings of 11th IEEE International Symposium on Multimedia*, pp. 477-82.

Wiggins, G. A. (2010). A cross-domain model? Grouping of phonemes into syllables by a model of melodic segmentation. In *Proceedings of ICMPC-11*.

Wiggins, G. A., Miranda, E., Smaill, A., & Harris, M. (1993). A framework for the evaluation of music representation systems. *Computer Music Journal*, *17*(3), 31-42. Machine Tongues series, number XVII.

Wiggins, G. A., Pearce, M. T., & Müllensiefen, D. (2009). Computational modelling of music cognition and musical creativity. In R. Dean (ed), *Oxford Handbook of Computer Music and Digital Sound Culture*. Oxford University Press.

Woods, W. A. (1970). Transition network grammars for natural language analysis. *Commun. ACM*, *13*(10), 591-606.

Zatorre, R. J. & Halpern, A. R. (2005). Mental concerts: Musical imagery and auditory cortex. *Neuron*, *47*(9-12).

• Sobre la inexistencia de la música: por qué la teoría de la música es una quimera de la imaginación

Abogamos por un acercamiento a la teoría de la música que parte de la posición de que la música es primariamente una construcción de las mentes humanas (y secundariamente una construcción social) y ello contrasta con el acercamiento implícito en el trabajo de algunos teóricos de la música, que tratan la música como si fuera un absoluto *cuasi* platónico definido externamente. Argumentamos que una conclusión natural de esta aproximación es que la teoría de la música, aunque ya es una especie de psicología folk, puede beneficiarse del hecho de ser informada más explícitamente por los estudios de cognición musical. Damos ejemplos del trabajo en modelos computacionales de cognición musical, siguiendo nuestro acercamiento, que intenta situar cada fenómeno musical en un contexto ecológico motivado por consideraciones evolutivas, y que intenta explicar fenómenos musicales independientemente de la intervención explícita del teórico. Argumentamos que sólo así puede un modelo ser expuesto verídicamente para explicar la fenomenología de la música. Situamos nuestro argumento en el contexto de la Teoría Generativa de la Música Tonal (Lerdahl y Jackendoff, 1983), las Lingüísticas Generativas, y otras aportaciones de este volumen, y comparamos todos ellos con los resultados de estudios de modelización basados en nuestro propio acercamiento.

• Sulla non esitenza della musica: perchè la teoria musicale è un'invenzione dell'immaginazione

Sosteniamo un approccio alla teoria musicale che parte dal principio che la musica sia innanzitutto una costruzione della mente umana (e in secondo luogo una costruzione sociale) confutato attraverso l'approccio, implicito nel lavoro di alcuni teorici musicali, che considera la musica come un assoluto quasi platonico definito esternamente. Sosteniamo che una conclusione naturale di questo approccio sia che la teoria musicale, già considerabile una forma di psicologia popolare, possa trarre beneficio dall'essere arricchita più esplicitamente dagli studi cognitivi sulla musica. Forniamo alcuni esempi dalle acquisizioni nel campo dei sistemi computazionali della cognizione musicale, seguendo un approccio il cui obiettivo è di collocare ciascun fenomeno musicale in un contesto ecologico motivato da considerazioni evolutive e che aspiri a spiegare i fenomeni musicali indipendentemente dall'intervento esplicito del teorico. Riteniamo che solo in questo modo un modello può considerarsi valido nell'esplicitazione della fenomenologia della musica. Collochiamo la nostra riflessione nel contesto della GTTM (Lerdahl and Jackendoff, 1983), della linguistica generativa e di altri saggi del presente volume comparandoli con i risultati di studi di sistemi basati sull'approccio da noi abbracciato.

• De la non-existence de la musique : pourquoi la théorie musicale est-elle un produit de l'imagination

Nous argumentons en faveur d'une approche à la théorie musicale qui part du point de vue selon lequel la musique est, avant tout, une construction de l'esprit humain (et secondairement une construction sociale). Nous contrastons cette approche avec celle, implicite dans l'œuvre de certains théoriciens de la musique, qui traite la musique comme si elle était un absolu quasi platonicien défini extérieurement. Nous avançons qu'une conclusion naturelle à cette approche est que la théorie musicale, bien qu'elle soit déjà une sorte de psychologie populaire, ne peut que gagner à être plus explicitement informée des études sur la cognition musicale. Nous donnons des exemples tirés des travaux sur la modélisation computationnelle de la cognition musicale, d'après notre approche qui tente de placer chaque phénomène musical dans un contexte écologique, motivé par les considérations évolutionnistes, et qui tend à expliquer les phénomènes musicaux indépendamment de l'intervention explicite du théoricien. Nous soutenons que seulement ainsi un modèle peut être déclaré conforme à la réalité pour expliquer la phénoménologie musicale. Nous situons notre argumentation dans le contexte de la théorie générative de la musique tonale (*GTTM*) (Lerdahl et Jackendoff, 1983), de la linguistique générative, et d'autres articles dans le présent volume, et les comparons avec les résultats des études de modélisation basées sur l'approche que nous avons adoptée.

• Über die Nicht-Existenz von Musik: Warum Musiktheorie ein Produkt unserer Einbildung ist

In diesem Beitrag gehen wir davon aus, dass Musik primär ein Konstrukt des menschlichen Geistes (sowie in zweiter Hinsicht ein soziales Konstrukt) ist und vergleichen diese Position mit der Perspektive einer Reihe musiktheoretischer Ansätze, wonach Musik ein externes quasi-platonisches Absolutes darstellt. Als natürliche Schlussfolgerung unseres Ansatzes halten wir fest, dass Musiktheorie — die an sich schon als eine Art Alltagspsychologie gelten kann — von Ergebnissen der Musikkognitionsforschung profitieren kann, wenn sie sich mit diesen explizit und ernsthaft auseinandersetzt. Beispiele aus dem Bereich der Computermodellierung von Musikkognitionprozessen sollen den Ansatz verdeutlichen. Bei der Modellierung dieser Prozesse wird versucht, das jeweilige musikalische Phänomen innerhalb eines ökologisch validen Kontexts zu betrachten, welcher evolutionäre Überlegungen miteinschließt und zusätzlich darauf ausgerichtet ist, musikalische Phänomene unabhängig vom direkten Eingriff eines Musiktehoretkers zu erklären. Wie wir darlegen, kann ein Modell nur unter diesen Premissen als ein Model gelten, das tatsächlich Musik phenomenologisch erklärt. Die Argumentation dieses Beitrags findet im Kontext der generativen Theorie tonaler Musik (Lerdahl and Jackendoff, 1983) statt sowie im Bereich der generativen Linguistik und auf dem Hintergrund der übrigen Beiträge in diesem Band, welche wir mit den Ergebnissen von Modellierungsstudien vergleichen, die unserem Ansatz folgen.