
John Halle
PART ONE. Whose (post)-Structuralism? – A Response to Pat McCrelles

1. Introduction

As the title suggests, part one of this paper was originally delivered as a Yale departmental talk in response to recent work by my colleague Pat McCrelless who was then attempting to incorporate into musical scholarship certain ideas and insights which had been developed under the general heading of post-structuralism. My main objective was to bring attention to the fact — and to the consequences of the fact — that there were really two distinct variants of structuralism, each provoking a distinct reaction. One of these was structural anthropology formulated by Claude Lévi-Strauss the repudiation of which by Kristeva, Lacan, Foucault and other highly influential figures constitutes what we now known as post-structuralism. Within a prior generation, however, a very different variant of structuralism was taken for granted, namely structural linguistics associated with

1. The following ruminations on the scope and limits of cognitive theories, the eclipse and possible resurgence of what I call cognitivist post structuralism, and the relevance of certain foundational debates within the field of linguistics to musical scholarship were written more than a decade ago. Signata editor Per Aage Brandt felt that they would be of interest to Signata readers and helpfully suggested combining what were originally two distinct essays into the version published here. (Those interested in the original versions can find them posted on my website here http://johnhalle.com/musical.writing.technical/structuralism.music.pdf and here http://johnhalle.com/musical.writing.technical/columbia.roundtable.pdf). I would like to extend my gratitude to Prof. Brandt for recovering this work and for his editorial supervision. Errors in fact and interpretation are, of course, my responsibility.
relatively unknown figures such as Leonard Bloomfield, Roman Jakobson, Bernard Bloch, and Edward Sapir. The scholarly tradition which derived from the critique of linguistic structuralism, namely, generative grammar, is not denoted by the term post-structuralism, though it is apparent that, as a matter of logic if not practice, it has as legitimate claim to that status as does the later tendency.

That this is the case is not widely understood. With the exception of Chomsky, those who have worked within the generativist paradigm, including, for example, Howard Lasnik, Ken Hale, Joan Bresnan are as little known as the structuralists themselves. This is largely understandable in that their work, like that of the structuralists, has dealt with technical subjects in syntactic and phonological theory. That it involved overturning assumptions of structuralism was implicit in the methodological approach with most of them tending to leave to others to explain why their rejection of core structuralist assumptions was conceptually and methodologically defensible. Many of the fundamental arguments were provided in seminal works by Chomsky (1976, 1980, 2000b) though recent work by Boeckx & Piattelli-Palmarini (2005) and Embick & Poeppel (2005) has also usefully re-addressed some of these questions.

The following uses this work as a window onto why the assumptions of structural linguistics, while reasonable, were wrong, though clearly, in retrospect, productively wrong. Specifically, it became apparent that the common sense view of the empirical basis of linguistic theory as consisting in the external manifestations of language needed to be rejected. What was required to replace it was the recognition that the essential reality of language is abstract, internal and psychological, its underlying nature only indirectly revealed by theoretical models based on the idealized intuitions of a fluent speaker.

Insofar as it has paid dividends for the study of language, the question is naturally raised whether this methodological perspective, what might be called cognitivist post-structuralism, can be usefully transferred to another human activity in which we exercise the capacities of our brains to make sense of what we hear and to create structures which are communicated to others, namely music. That this is the case will be made apparent by sketching some of the details of Fred Lerdahl’s model of pitch representation, an abstract model which captures significant aspects of the unconscious knowledge experienced listeners exercise when they make sense of the music they hear.

While recognizing the importance of the work within this tradition, it is clear by now that cognitivist post-structuralism did not amount to a broad scholarly tendency, as did the continental post-structuralism which McCrelless among others was then attempting to incorporate into musicology. That said, in recent years certain shortcomings of post-structuralism have begun to be recognized and taken as a basis for raising questions about the post-structural icons mentioned above.

Among these is Daniel Zamora’s (2014) critique of Michel Foucault as well as Eric Drott’s (2015) critique of Jacques Attali’s Noise, a post structuralist text highly
influential on the “New Musicology” of the 1980s and 1990s. Within this climate, a reconsideration and possibly a resuscitation of cognitivist post-structuralism seems possible. If this is to occur, however, there will need to be a general recognition that the insights it conveys are limited by constraints imposed by the scientific method. These necessarily entail that much of what we might want to know about any “natural object”, whether it is language, music or a planet, will be, in Hume’s words cloaked in “that obscurity in which they ever did and ever will remain.” (cited in Chomsky 2000a). Real, as opposed to illusory, progress may require that we are content with this state of affairs.

2. Structuralisms and music

The recognition of the limits imposed on our capacity to describe the facts of language has its roots in a fundamental shift in how linguistics viewed the central object of its study in the mid 1950s. Prior to that point, it was generally believed that a complete description of all the world’s languages was within our grasp, as techniques for mechanized recording and transcription were developed. For that reason, it would probably have come as a surprise to psychologists, anthropologists, and linguists that the then dominant tendency in the field, structuralism, would, in the space of only a few years, become a moribund intellectual tradition. Not long after, however, even the most committed structuralist would have been hard pressed to claim otherwise and by now most of the leading names of structuralism are, as mentioned above, largely unrecognized.

Within musical scholarship, while post-structuralism has exerted a persistent and significant influence, as Pat McCrelles pointed out some years back (McCrelles 2000), a renewed interest in structuralism may have been provoked by Hatten’s (1994), Agawu’s (1991) and Nattiez’ (1990) investigations into musical semiotics, a field which McCrelles refers to as “the quintessential offspring of structuralism.” But semiotics, I will argue, while certainly an influential offspring of structuralism, should not be seen as typical. For unlike semiotics, which remains an active and prolific research paradigm, traditional structuralism is primarily known for its having failed. This is particularly the case within linguistics—the discipline which originated structuralism and provided the conceptual and taxonomic framework subsequently imported to other humanistic disciplines.

That it is primarily linguistics that has been able to learn from the collapse of structuralism is the subject of Leonard Jackson’s (1991) polemic The Poverty of Structuralism. Part of the reason this is the case, Jackson argues, results from the fact that the variant of post-structuralism which has become dominant in the humanities and social sciences (but not in linguistics) has manifested a particular myopia with respect to structuralism. That is, it tends to identify structuralism narrowly with the structural anthropology of Levi-Strauss and his continental followers on one side, and the work of Ferdinand de Saussure on the other. That structuralism had
a continuous and active existence for the two intellectual generations between
Saussure and Levi-Strauss tends to be ignored. As a result there is comparatively
little awareness of the rich-albeit somewhat technical work from the first half of the
20th century which Saussure’s strictly linguistic analyses inspired.

Jackson refers to “the great misreading of Saussure” which has emerged from
post-structuralism, in which:

Every theoretical distinction … is taken up, enlarged upon, analysed and
transferred into other contexts in which it often means something entirely
different. But the linguistic examples that pin down these theoretical distinctions
are usually ignored. The (text) is read; but the italic type… indicat(ing) linguistic
examples … is skipped over. (Jackson, p. 104-105)

Terms which began as a precise technical vocabulary in Saussure and subsequent
variants of structuralism would be imported to a range of other disciplines. Most
notable among these would be Levi-Strauss’s structural anthropology based
on terminology derived by Roman Jakobson such as the marked-unmarked
dichotomy, the distinctive feature, binary opposition etc. While these analytical
categories surely provided important insights into cultural and literary practice,
much precision and, more importantly, many rigorous, empirically grounded
conclusions having to do with one human activity, namely language, were lost as
the linguistic roots of structuralism became obscured. I will suggest that it is only
when structuralist concepts are returned to the initial context in which they were
advanced that the real lessons of structuralism’s fall can be learned. In order to do
so we need to make a brief attempt to deal with the italic type which Jackson refers
to the facts of language as they were understood by Saussure and other scholars
from the first half of the twentieth century. In so doing it will also become apparent
why the structuralist framework within linguistics was ultimately found to be
scientifically unsatisfactory: it simply could not account for these facts.

2.1. First Generation Structuralism

The idea that “facts” of particular languages divorced from the historical continuum
of language could be the object of study, while uncontroversial today, was by no
means so when linguistic structuralism began to emerge in the nineteenth century.
19th century linguistics as practiced by the neogrammarian school of Herman Paul,
Karl Bruggman and also the first American linguist of international reputation,
Yale Professor William Dwight Whitney, had been focused on an explanation of
linguistic form in reference to the historical development of languages. According
to Anderson (1985), the “historical study of language was the only genuinely

---

2. For example, the names Leonard Bloomfield, Zellig Harris, Franz Boas, and Edward Sapir,
account for fewer than ten citations in Dosse’s monumental history of Structuralism (Dosse,
a similar distribution.
scientific approach to the facts of language” and formed the basis of linguistic scholarship. In this respect, 19th century linguistics bears comparison with musical scholarship which would develop its own “synchronic” discipline, music theory, only after the diachronic field of musical scholarship, musicology, had been long established. Much discussion outside of linguistics has been devoted to Saussure’s “diachronic” versus “synchronic” distinction, but it should be understood that at least one basic aspect of the distinction was directed internally. For a synchronic approach to linguistic analysis was to, in Anderson’s words, “completely change the direction of the field”. In so doing, Saussure invented the field of “descriptive” or “structural” linguistics. Just as the field of music theory allowed for an exclusive focus on the formal richness of particular works, the synchronic study of language would reveal that particular languages were tremendously rich and varied. The picture would be very different from the orderly transition of language to language revealed, as it were, from on high in the laws of historical sound change devised by Grimm and Verner, the cornerstones of the neogrammarian linguistics. Indeed, actual existing languages when looked at from a structuralist perspective were a veritable riot of sounds and physiological gymnastics. The recognition of the daunting complexity of all languages was advanced by results from the ambitious field research programs undertaken by structuralism. Many of these were studies of languages of indigenous peoples having only the most distant relationship to the indo-European languages which formed the basis of the neo-grammatician canon. Rather than relegating “exotic” languages to the periphery, structural linguistics drew scholarly attention to these languages as a validation of the inherent interest in synchronic inquiry. What emerged from structuralist fieldwork — precise descriptive grammars of African click languages, indigenous Australian languages without apparent word order, languages without syllable structure, languages in which parts of speech were interchangeable — would constitute much of the empirical domain which linguistic theories to this day are required to account for.

2.2. The Abstract Phoneme

Structuralist fieldwork was undertaken as part of a larger ethnographic program including in addition to linguistics, anthropology and sociology, what was then known as “folklore”, later developing into what is now known as ethnomusicology. Whereas folklorists had at their disposal a notational system which could specify to a reasonable degree of precision the objective structure of music, it would not be until well into the 19th century that linguistics would develop a notational system of roughly equal descriptive power. This was the phonetic alphabet receiving its exposition in A.M. Bell’s Visible Speech of 1867.

The phonetic alphabet derived from a theoretical advance in 19th century linguistics which posited that languages are defined by a limited repertoire of speech sounds known initially as phones and later as phonemes. These phones can in turn be seen as composites of what would be known as “features”: discrete
physiological gestures effected by the “speech organs” including the lips, tongue, mouth and throat. The phonetic conception of language adopted by structuralism was fundamentally an atomistic one and as such was seen as a breakthrough equivalent to what molecular theory had been for the natural sciences a century earlier. Bell’s alphabet, and later developments of it, would function as a kind of periodic table whose finite means allowed for the description of the practical infinity of the world’s languages.

While the phoneme is relatively familiar, the linguistic issues surrounding its status as a verifiable element of actual language are not. For just as the physical atom would give rise to centuries of speculation as to its ultimate existence so would the notion of the phoneme as a “linguistic atom” raise as many questions as it answered. One such question became immediately apparent when in cataloguing the speech sounds of a language a clear choice between lumping and splitting confronted linguists attempting to catalogue the phonemes of languages. For example, English, like most languages, makes use of the unvoiced dental consonant “t” formed by the closing off of the air cavity by the tongue behind the teeth. In certain contexts — as the initial consonant in the word “tea”, this consonant is accompanied by a burst of air known as aspiration. As the final consonant of “scat”, however, it will be noticed that aspiration disappears (at least in the speech of native speakers). When “t” appears as the medial consonant of “writing” no actual complete constriction of the airflow takes place, and the stop is replaced by the “tap” or “flap” in which the tip of the tongue briefly touches the roof of the mouth immediately behind the teeth.

Since three distinct physiological actions are implicated in the production of this sound, an objective “phonetic” description of the external structure of these utterances must register the appearances of at least three distinct composites of features. An “objective” description of the consonants of English will be seen to require not just the three variants (or allophones) of the English “t” mentioned above but also numerous other intermediate articulatory gestures, including partial flapping, degrees of aspiration, half-voicing, etc. Furthermore, a complete specification of the physiological or acoustical structure of actual speech would require English grammars to represent intonational contour, tongue clicks, and even hand gestures, phonetic characteristics which are highly salient or “distinctive” in the phonological structure of other languages but are not in English. A theory which claims to describe the objective reality of language cannot, by definition, arbitrarily choose to ignore that which the analyst or the speaker subjectively “knows” not to be significant. But to not allow for such intuitive judgments as to what constitutes the relevant facts of language was to invite chaos. What Anderson refers to as the “fully specified surface” was as much a practical impossibility as it was a theoretical dead end.

Not only would a complete specification of external speech be unwieldy, more crucially it would fail to capture a more fundamental fact about the perception
of the sound structure of language also noted by Sapir. Namely, that “the actual rumble of speech” is perceived as “an ideal flow of phonetic units.” Speakers of a language understand that:

Significant structure of language is dependent upon the unconscious selection of a fixed number of “phonetic stations” or sound units. These are essentially modifiable … but the essential point is that … definite psychological barriers are erected between phonetic stations so that speech ceases to be an expressive flow of sound and becomes a symbolic composition with limited materials. (Sapir, 1933, p. 58. n. 25)

In other words, while the external reality of the units of speech is highly continuous, the internal, psychological reality of speech is discrete. The smallest units, the phonemes, will be “modified”, sometimes quite radically in actual speech to embody the range of allophones which are observed in external language. But no matter how much they are distorted, all instances of the phonetic realization will result in the speaker decoding and recovering the underlying phoneme. While the compilers of structuralist grammars were generally not entirely explicit on this point, they would reflect speakers’ “unconscious” understanding of the “psychological barrier erected between phonetic stations” (p. 8) by representing subjectively similar but objectively highly disparate phonetic units as belonging to the same phonemic categories.

2.3. Music and Language: A Structuralist Comparison

Sapir, who studied composition under MacDowell at Columbia, would draw a useful analogy directly relevant to musical scholarship. The status of phonetic versus phonemic structure in language is, Sapir contends, comparable to the listener’s experience of music:

The analogy (of phonology) with musical theory seems quite fair. Even the most resplendent and dynamic symphony is built up of tangibly distinct musical entities or notes which, in the physical world, flow into each other in an indefinite continuum but which, in the world of aesthetic composition and appreciation, are definitely bounded off against each other, so that they may enter into an intricate mathematics of significant relationships.” (Sapir, 1933, p. 8)

The phoneme is a kind of psychological phantom which a fluent speaker is able to project onto a linguistic surface. The surface itself, however, provides very limited and impoverished clues from which listeners construct a rich phonemic interpretation. The “significant relationships” of music, Sapir claims, whether these are harmonic, melodic or rhythmic, are equally mental constructs: not components of the piece per se but psychological categories projected by the listener onto a surface structure of music which allows for a practically infinite class of interpretations.
It is not unreasonable to assume that Sapir, who would argue for the inherent “psychological reality of the phoneme” (the title of one of his best known essays [Sapir, 1933]), would accept that ultimate reality of the work of music was also similarly psychological. Unfortunately, Sapir’s early death in 1939 prevented his developing the ideas suggested in the above passage further. (There are, however, tantalizing hints in Sapir’s 1921 essay “The Musical Foundations of Verse.”) It would be some time before this position could be articulated directly as it would be by Ray Jackendoff (1992):

> What is the “objective reality” of the *Eroica Symphony* or *Stardust*? It seems to me that these pieces are best considered to be abstractions that can be realized in performances or scores of wildly varying physical characteristics (it is possible, though appalling, to conceive of the *Eroica* played by a quartet of marimbas; and of course a popular song like *Stardust* can be subjected to endless arrangements). But the abstraction that unifies different performances of a piece of music is precisely captured in the mental representation of music — the mental structures that performances evoke. That is, the constancy and reality of a piece of music are purely mental. (Jackendoff, 1992, p. 165)

Jackendoff’s “internalist” perspective on music would and, indeed, could probably only have come after generative linguistics had established Sapir’s belief in the ultimate “psychological reality” of the units of language as conventional wisdom — at least within linguistics. During his lifetime, Sapir’s views would be derided by most structuralists who, highly influenced by empiricist and behaviorist tendencies afoot at the time regarded such “mentalist” explanations as unscientific or worse.

With the mind off limits, it was necessary for structuralists to posit the phoneme as existing “out there” in the acoustical surface of language. While Bloomfield and others recognized that they were making a significant leap beyond the empirical data in positing a verifiable acoustic basis for the relationship between the phoneme and its allophones they were hopeful that it would emerge.

Unfortunately, rather than confirming structuralist expectations, as acoustics developed it failed to provide the objective basis for these or other seemingly subjectively obvious linguistic relationships. The central phonemic reality of linguistic structure would be, like that of the atom itself, not verifiable through direct observation. Rather both entities, the atom and the phoneme, were theoretically required in order to account for the range of phenomena which are empirically observed. The phoneme was a psychological phantom but a necessary one for any reasonable account of linguistic structure to be advanced.

---

3. Bloomfield (1933) is explicit on this point: the phonemes, according to him, are specific “features of sounds which speakers have been trained to produce and recognize in the current of actual speech sound.” (my italics). Alternatively, in Bloch’s (1941) formulation of Bloomfield’s position, the phoneme defines the “constant feature of the soundwave” or “the central member of a class” of sound waves around which the allophones form a statistical “cluster”. (Bloch, 1941, p. 93).
2.4. The Collapse of Structuralism

It is generally agreed that the demise of structuralism can be dated to the rise of generative grammar initiated by Noam Chomsky’s work of the early 1950s. While not all of the core theoretical assumptions of generative linguistics would be made explicit in early work in the field, linguists would more or less tacitly assume a strong version of Sapir’s argument for the “psychological reality” not just of the phoneme but of all significant linguistic structure. As the field matured, it would come to accept Chomsky’s formulation that i-language, the internal, intensional, intuitive aspects of language would form “the object of inquiry” for linguistic science (Chomsky, 1986). The “facts” of e-language, the external artifacts of language which defined structuralist inquiry, would be relevant to linguistics only insofar as they were relevant (either positively or negatively) to theories of the mental representation of language. Linguists would shift their orientation from studying the structure of language as it is acoustically or physically embodied in the real world to the psychological mechanisms which underlie linguistic competence. Or, as put somewhat differently by Mark Baker (2001), since Chomsky linguists study not language as such but rather mental recipes which speakers appeal to when they produce or consume language. A more complete overview of contemporary linguistics is beyond the scope of what is necessary for our purposes here. For those interested, I would recommend, in addition to Baker’s excellent The Atoms of Language, Anderson and Lightfoot’s The Language Organ (Anderson and Lightfoot, 2002) and, somewhat more provisionally, Steven Pinker’s The Language Instinct (Pinker, 1994). These texts go some way towards correcting the dearth in non-technical introductions to linguistics, making accessible to those outside the field what lies behind the formal discourse of contemporary work.

An unfortunate consequence of the absence of readable non-technical introductions to the field is that there is little awareness outside the field of what linguists work on and why. It is maybe understood by non-specialists that, in a general sense, linguists attempt to derive rules which describe grammatical constructions. Beyond this, little is understood about how these rules work and still less is known of the remarkable success which linguists have had in discovering general principles — or, more specifically, parameters — which determine the underlying structure of a wide variety of superficially disparate languages.

4. Chomsky (2000b) describes the shift from structuralism to the generativist (or cognitivist) model in the following way: “There was an important change of perspective: from the study of behavior and its products (such as texts), to the inner mechanisms that enter into thought and action. The cognitive perspective regards behavior and its products not as the object of inquiry, but as data that may provide evidence about the inner mechanisms of mind and the ways these mechanisms operate in executing actions and interpreting experience. The properties and patterns that were the focus of attention in structural linguistics find their place, but as phenomena to be explained along with innumerable others, in terms of the inner mechanisms that generate expressions”.


More importantly still less is known of real, if only partial, answers, grounded not in speculation, but in “hard” scientific theory, which are now emerging on a range of questions which, until now, could barely be formulated coherently let alone answered. Among these are: How is language constructed so that we can understand it? How does language relate to thought and to what extent does it exist independently of thought? What aspects of language are learned, and which are innate?

Progress on these and other longstanding, complex questions has been made possible by linguists having correctly identified the essential object of study of their field as the underlying mental representation of language, not the external artifacts of language, or language “out there”. In other words, the progress of linguistic science is directly related to its having overthrown the basic assumptions of structuralism. Whether a post-structuralist\(^5\) music scholarship promises to obtain equally substantial results is, I recognize, a controversial position, a point I will return to. For now, I will argue that developing a poststructuralist or, to use Chomsky’s terminology, cognitive perspective should be a higher priority within musical scholarship than it is at present.

I say this for two reasons. First, as just noted, linguistics has demonstrated that the potential upside is considerable — real answers to questions which have been obscure emerged once the essential mental character of language was recognized as the “object of inquiry”. Among the questions which might emerge from obscurity from within a post-structuralist perspective on music are — what is large scale musical form, to what extent is it an objective characteristic imposed by composers on compositional materials, and to what extent is it a psychological primitive required of all coherent musical structure in the same way that, for example, all coherent linguistic utterances are constructed from linguistic primitives such as phonemes, lexical categories and phrase structures. A second question, one which has been at the root of much heated debate in new music circles, is to what extent are certain compositional systems, to use Lerdahl’s useful terminology “cognitively opaque” (Lerdahl, 1992)? To what extent is it required for listeners to “assign works of music a detailed representation” and what particular compositional structures allow for this psychological process to be enacted. And what sorts of structures are resistant to processing, even given practically unlimited training. This is not to say that all, or even many, of the questions we would most like to have answered would be addressed within a cognitive framework. However, some, I believe, such as those posed above might be. And this prospect should be sufficiently attractive for some room to be cleared within musical scholarship for the “cognitive perspective”.

---

5. I recognize, of course, that the term has come to denote various post-modernist approaches in literary criticism, anthropology and cultural studies. I do not believe that this appropriation is legitimate, for reasons which should be clear. In any case, I trust that by this point, my use of the term will not cause confusion.
Second, a cognitive approach to musical scholarship would seem to hold out the hope for similar successes to those achieved in linguistics because, while there are clearly enormous differences, both are at root essentially cognitive activities. Both music and language can be manifested externally in our lips, tongues, fingers, on recordings and on the printed page. But underlying these behavioral artifacts is the existence of music in our minds, specifically as a mental representation. If this is so, there is good reason to study it as such. The study of I-language was found to be a practical necessity for linguistic science, and it is likely that the analogous conceptual and methodological category of I-music should become at least relatively familiar to those in musical scholarship.

Given the both potential upside and the fundamental common ground, it is reasonable to expect music scholarship to have a better grasp of the underlying issues surrounding linguistic structuralism — what it is, and what it is not — than it generally has displayed. Indications of lapses on this score are, unfortunately, relatively easy to discover. Among these are Ian Bent’s New Grove’s entry on structuralism (Bent, 1980) in which Chomsky is identified as “a chief exponent of structuralism”. A more recent instance is found in an online contretemps between Joe Dubiel and Allen Forte where Forte refers to “the structuralist linguistics of the seventies”. (One assumes that Forte is referring to various attempts to shoehorn then fashionable ideas of Chomsky into music theory). I hope that at minimum, the above discussion makes clear that this view is incorrect: to reiterate, not only is Chomskyan generative grammar a direct repudiation of structuralism, the successes of contemporary linguistics were directly linked to its explicit rejection of the fundamental premises of structuralism.

3. Two loose ends

I will close this part by making two qualifications to claims I have advanced in the preceding. The first is to note that the paths of generative linguistics and music theory have, of course, crossed in the past. Bernstein’s 1975 Norton Lectures published as “The Unanswered Question” was an early, albeit problematic, attempt to graft what was then called transformational grammar onto musical analysis. Lerdahl and Jackendoff (1983) and to a lesser extent Narmour (1990) have offered theories of musical representation in some ways consistent with the assumptions of what might be called Chomskyan post-Structuralism. I want to suggest here (and I will argue elsewhere) that it is not always entirely clear that the implication-realization model or the generative theory is consistent with the larger picture of the mind which Chomsky’s work strongly argues for. Indeed, as what has been

---
6. The confusion of Forte and Bent is understandable. (See note 5). Furthermore, most post-structuralists, do, in fact, categorize Chomsky not as a post-structuralist but as adherent of structuralism (see Dosse 2001).

7. Chomsky would, of course, vehemently object to the term. See Chomsky (1968).
called “cognitive music theory” has developed, it has tended increasingly to embrace assumptions more consistent with structuralism than with generative linguistics. Most notable among these is the expanding field of music perception which is largely concerned with empirical studies of musical behavior. These tend to narrow focus on the “objective” external facts of musical behavior (as opposed to the psychological mechanisms underlying the behavior). Thus, much work in music cognition is highly reminiscent of the more rigid variants of structuralism—even behaviorism. Structuralist assumptions have also, surprisingly, found their way back into that which was initially inspired by generative grammar. Lerdahl’s pitch space “journeys” (Lerdahl, 2001) can, with some difficulty, be construed as they are evidently intended, that is, as psychological representations. However, as can be seen in, for example, Gollin (2001), they are just as easily understood as traditional structural analyses concerned with highlighting the composed-in characteristics of the tonal structure of the particular work under analysis and only secondarily with modeling a plausible mental representation of harmonic relationships.

The second qualification is to note that a Chomskyan view of music as a purely mental “natural object” is inherently problematic in certain important respects. Clearly, the objective structure of musical works, as opposed to its psychological representation, can, and in the case of notated music certainly does, exist as an autonomous entity, independent of, and on occasion impervious to, the perceptions of any listener, experienced or otherwise. Since the Middle Ages, composers have created complex and elegant relationships within pieces or, more precisely, secret codes at least some of which are not “heard” in any meaningful sense, but are only observable by those who have been able to decipher the underlying generative system. Then there is a significant gray area of works in which certain composed-in features of music can become perceptually accessible, either through practice or familiarity with the compositional practice which gave rise to them. In short, what we hear (or perceive) can be influenced by what we know about the underlying organization of the perceptual stimulus and, as argued by Fodor (1988) and Churchland (1984), “perceptual plasticity” needs to be incorporated into any reasonable theory of mental structure.

Given this essential fact of musical experience, it is entirely reasonable for music theory to attempt to discover and categorize the underlying structure of works whether this structure is immediately and unproblematically perceptually salient, provisionally perceptually salient or completely impervious to our perceptual faculties. Furthermore, as a pedagogical matter, it may be eminently appropriate to view our perceptual faculties as highly contingent and trainable indeed, close to “infinitely malleable” in a sense that Chomsky would find uncongenial, if not outright abhorrent. Art music, as an aesthetic discipline as opposed to a communicative system, requires setting the bar high for those who want to fully

---

8. See, for example, Chomsky (1968).
experience it. This does not mean, of course, that we should engage in the futile task of attempting to internalize musical structure which is, to appeal again to Fred Lerdahl’s useful term, “cognitively opaque”. It does mean that we should not be willing to give ourselves, or our students, the easy way out in assuming, as we do with language, that everything which is there to be perceived is designed so as to be readily accessible.

But to have a range of options for theoretical inquiry available to us also means that certain directions are likely to be more fruitful and productive than others. The lessons of the collapse of structuralism for linguistics demonstrated precisely that. Namely, an investigation into the external structure of language, while producing much important empirical data and many interesting observations, was ultimately sterile in that it was incapable of supporting an empirical theory of language. Based on this history, a reasonable approach to the long standing question “what is a theory of music a theory of?” needs to involve a full awareness of the essential psychological character of music and recognize that the external artifacts of music — music as heard in performance or as notated in scores — are the shadows cast by the underlying psychological reality. The legacy of structural linguistics and its honorable failure provides us a clear precedent for how a successful inquiry along these lines can proceed.


1. Metrics

The following discussion is focused on one approach within what I characterized above as cognitivist post-structuralism, namely, Fred Lerdahl and Ray Jackendoff’s Generative Theory of Tonal Music and its extensions within Lerdahl’s Tonal Pitch Space and elsewhere (I will refer to all of this work hereafter as GTTM). Insofar as this body of work has succeeded, it has, of course, provided certain answers. But perhaps more importantly, the systems of representation advanced within it have allowed us to “see” the phenomena in question more clearly allowing us to ask questions having to do with what it means to “know” a piece of music. In particular GTTM does so by breaking the question down into component parts. Exemplary in this respect is the aspect of Lerdahl’s theory having to do with our knowledge of the inferred rhythmic structure of music and the formalism which must be posited in order to account for it. After having been first advanced three decades ago, GTTM’s descriptive apparatus for representing the intuitions implicated in metrical structure, namely, the grid representation shown immediately below applied to the Haydn minuet in example 1 has gained general acceptance.

The acceptance of this formalism, like that of any other formalism, resides crucially in its capacity to capture important characteristics of the facts under consideration thereby allowing us to view the essential mental reality of metrical intuitions obscured within other representations of inferred metrical structure. Thus, to take one such example, the prosodic representation of metrical structure proposed by Cooper and Meyer shown in Example 2 is a reasonable attempt at characterizing metrical intuitions.

The essential shortcoming inherent in the prosodic representation, as noted in GTTM, is in its attribution of metrical prominence to portions of phrase segments. This can be seen by observing the markings within the horizontal brackets in example 2. It will be noticed that the two eighth note upbeat and the entire first measure of the passage is assigned a macron on level 2 by Cooper and Meyer. This is to indicate that this half of the bracketed time span is relatively strong compared to the second half of the bracket which, as indicated by the breve, is heard as weak. Metrical prominence is, however, understood by listeners as an inherent characteristic not of continuous segments but of discrete points in time, that is, of the temporal locations marked by what we refer to as the beat. Thus, it is the first beat of the first measure, and not the time span extending from the upbeat to the third beat of the measure which is heard as strong. The grid representation identifies the former point in time as strong by means of the five dots assigned to this precise location while the latter is identified as relatively weak by virtue of its being assigned only a single dot. Since the grid correctly represents these and other
facts\textsuperscript{10} which are essential components of our understanding of meter, it and not the prosodic representation is now broadly adopted, including among those who are not necessarily sympathetic with the cognitivist approach in explaining the metrical structure of music.

Another contribution of GTTM to our understanding of the mental representation of rhythmic structure is the recognition that the “accent” associated with the awareness of meter is of a special perceptual type, namely the metrical accent, a perception of relative prominence as an inherent characteristic of specifiable temporal positions. The metrical accent is categorically distinct from what is informally understood as an “accent” namely, the phenomenal accent, which is inherent in the physical signal itself resulting from a local extreme in amplitude, pitch, or sudden alteration in timbre. While it is natural for listeners, even experienced ones, to conflate all types of accent, they are distinct cognitive objects: not all metrically strong positions correspond with a phenomenal accents, nor are phenomenal accents enjoined from occurring on relatively weak positions of the metrical hierarchy. The phenomenal accent can function as an important cue to metrical accent; however, it is only one factor within a complex computation and is not determinative of the result. While they are easily confused, any reasonable account of musical structure must distinguish between these fundamentally distinct types of accent, and GTTM, while not the first to do so, was the first to provide a comprehensive and coherent account of their distinct character.

2. Epistemology

A significant achievement of GTTM’s approach to meter was the recognition that the informal term accent is in fact a kind of homonym, referring to two fundamentally distinct albeit interactive psychological entities. In order to make sense of metrical phenomena, GTTM made use of a gambit familiar within the natural sciences namely to devise a specialized terminology which overlaps with but which is not coextensive with informal usage. Just as the terms, work, energy, momentum, etc. have a precise, technical definition within physical sciences while evincing only a metaphorical connection to their use in such expressions as “I don’t have the energy to read that book,” “Joe Lieberman’s campaign is picking up momentum” and “Women’s work is never done,” the operative definition of “accent” within the GTTM system, as has just been discussed, is one which has only an oblique relationship to how we informally understand the term. While the folk science understanding may provide us with certain types of evidence with respect to our perception of meter in some instances, there is no reason why a theoretical account of the underlying basis of metrical structure and its computation should

\textsuperscript{10} For example, the grid will also be seen to be preferable in that it is able to represent the downbeat of the second measure as relatively strong compared to its neighbors but relatively weak compared to the first beat of the first measure.
be expected to conform to our informal sense of the term. Indeed to do so would guarantee its failure.

Insofar as GTTM is correct in the specific case of its account of metrical structure and its account of musical structure generally, knowing a piece of music means having access to the sorts of representations which are posited in the system, maybe not, indeed certainly not, exactly like them but having the broad outlines proposed therein: metrical structure designated by the grid generated by the beat induction computation outline therein or perhaps making use of proposals elsewhere,¹¹ constituency structure according to the computations designated by the grouping well formedness and preference rules,¹² vertical/harmonic structure within the outlines designated by prolongational and time span hierarchies. To the extent to which a listener has succeeded in constructing these representations, it is in this precise sense that one can be said to “know” a piece. More pointedly, insofar as a piece resists the construction of these representations, it can be said to be fundamentally unknowable, and perhaps, hardly a piece of music at all, a claim which has generated some controversy in the two decades since it was first proposed.¹³

To reiterate, these and many other insights with respect to the nature of musical form as it is psychologically represented constitute a contribution to our understanding of one aspect of what human knowledge is. But if one grants this to be the case, certain contradictions and confusions along the lines just mentioned immediately come to the fore not just with respect to musical knowledge but to knowledge generally. As we saw in the case of “accent”, just as the informal usage of the term obscures what is required in order to devise a descriptively adequate account of the phenomena, so too does the common usage of the term “knowledge” and the verb “to know” in many cases come into conflict with the technical sense of “knowledge” which the cognitive sciences require us to adopt.

It follows from this that it is likely that no generalized definition of the term will come close to offering an adequate account for all our uses of it. There is no such thing as knowledge but rather a menagerie of cognitive states of distinct types which we refer to by the same word but which are by no means the same thing. That this is the case seems fairly obvious. Thus, while we might say that

(a) “John knows the Eroica symphony” and
(b) “John knows that the moon is made of cheese”

— it is clear that the knowledge implicated in (a) relies on the sorts of computations and formalisms which form the basis of cognitive musical theory. As for (b), while this aspect of John’s knowledge must have an ultimate cognitive basis, there is no reason to believe a priori that the sorts of primitives, formalisms and computations which are appropriate to it are at all comparable to those which are

¹¹ See, for example, Eck (2001), Desain and Honing (1999), Povel and Essens (1985).
¹² Revisions to the grouping preference rules are proposed in Frankland and Cohen (2004).
appealed to in forming the sort of a musical representation which constitutes the musical knowledge implicated in the former statement. Although we use the term knowledge to refer to John’s relationship to both (a) and (b), there is no reason to assume that the term designates anything close to the same entity.

Examining the two sentences more carefully, whereas (b) necessarily implicates propositional knowledge on the part of John, his awareness of things in the world, their real world characteristics and their logically necessary and logically possible relationships, (a) assumes encoded psychological representations of the general variety proposed in GTTM. That there is an inherent ambiguity to the term we are employing is, of course, a common linguistic circumstance and does not necessarily pose any serious impediment to effective communication. The problem for us as investigators comes when we assume that the particular form of knowledge implicated in (b), namely propositional knowledge, is in some sense theoretically primary and on this ground must dictate the form of what we posit to make sense of our psychological capacities.

Why this matters for GTTM or, for that matter, any minimally plausible cognitive theory, is that it appears that the sorts of representations which one must posit within a cognitive theory of music are, it would seem, notable precisely for their absence of propositional content. A strong statement of this position can be found in Ray Jackendoff’s essay “The Problem of Reality” (1992). According to Jackendoff:

It hardly makes sense to say that the representations one constructs in response to hearing a performance of the Eroica (Symphony by Beethoven) are true or false. Nor does it make sense to claim one has propositional attitudes towards musical representations, which aren’t, as far as I can tell, propositions. The whole story becomes even more absurd if we think of the process of composition. Presumably, Beethoven wrote the symphony by virtue of creating mental representations (imagining music) that he wanted the written notes to evoke—in the absence of any overt musical signal. Were his representations false until he wrote the symphony down? Did they suddenly acquire a truth-value when the score was written or the first performance took place? I can’t make any sense of such questions—they entirely misconstrue the character of musical experience.

Musical representations provide a specific application of a general characteristic of cognitive representations which Jackendoff describes as follows:

A representation is not about anything. It does not strictly speaking represent anything, hence my hesitation about using it except as a … designator of what cognitive scientists believe the mind has in it.

If mental representations generally and musical representations specifically do not embody propositional attitudes, intentional states, or intentional content, what is a representation and what are its salient characteristics? According to Jackendoff, for the purposes of cognitive inquiry, a mental representation is “a
space of possibilities having an inherent organization.” Accordingly, a particular mental representation R

can be thought of as a location or region in the total space of possibilities. What makes R what it is is simply its relationship to other possibilities in the space — how it is distinct from other possibilities, what possibilities are close to it in what dimensions, and so forth. When we as theorists use symbols to state a theory of mental representation, it is not the symbols themselves that are significant but rather the range of distinctions possible in the system of symbols we adopt: these distinctions are claimed to be homologous to the organization of the relevant subsystem of brain states.

It should be apparent to those who are familiar with it that Lerdahl’s attempt to describe our perceptual relationship to the vertical/harmonic structure of tonal works provides a parade example of what Jackendoff has in mind in this passage and it is worth making a brief digression here to focus on certain specifics of Lerdahl’s model for this point to become clear. In particular, it will be noticed that the geometrical or more specifically topological organization proposed by Lerdahl in his Tonal Pitch Space has an “inherent organization” of the sort Jackendoff refers to which has the potential to express a wide range of logically possible perceptual relationships to harmonic structure. But only a small subset of these organizations are reasonable to entertain as cognitive models consistent with the facts.

3. Tonal relationship: pitch space

For this reason, it is appropriate to characterize Lerdahl’s investigation in Jackendoff’s terms as attempting to determine which of “the possibilities in the space” are optimal as a model for how we construct tonal relationships in pieces of music. Among these possibilities is the pitch space model which Lerdahl applies to the Chopin E major Prelude and which we repeat here as example 3 with certain alterations.

Example 3 fills out slightly more of the partial space represented in Lerdahl’s figure by identifying a few additional vertical and horizontal regions. The purpose
in doing so is to show one aspect of the “inherent organization” of the space more clearly, namely, while the horizontal axis is derived from alternating minor and major key regions a minor third apart, the vertical axis maps out major or minor key regions along the circle of fifths.

While this spatial orientation offers an attractive representation of the modulatory schemes of certain works, including the Chopin Prelude under consideration, the purpose of structuring the space in this way is not merely aesthetic, it is empirical. That is, it makes correct predictions with respect to the relevant empirical domain, which is, to reiterate, the listener’s unconscious understanding of one component of musical structure, namely the relative distance of tonal regions.

The central intuition which must be represented within a pitch space model is the perception of relative tonal distance of key areas or tonal regions, as we will refer to them: certain regions, namely, the dominant (V), subdominant (IV), parallel minor (i) and relative minor (vi), are understood as closest to the central tonic I. Others, for example, major keys a tritone or a minor second away from the tonic (Bb and F) are highly distant. As for the major third cycle of modulations characteristic of later 19th century harmonic practice and found in the Chopin prelude, these are somewhere in between: more distant and more surprising than the closely related keys though not maximally distant and unexpected. It will be seen that Lerdahl’s pitch space model succeeds in representing all of these regional distances accurately: the four most closely related keys are located immediately adjacent to the tonic I. Somewhat further away on the space are the key areas reached in the Chopin prelude. More distant still — requiring a traversal of two vertical and two horizontal steps is the bII region F major, and maximally distant within this model is the region built a tritone above the tonic Bb major which requires three vertical and two horizontal steps.

It should be obvious that not all models of pitch space which we might construct represent these and other regional distances accurately. This is what Jackendoff has in mind in referring to cognitive representations as defined by their “relationship to other possibilities in the space.” The relative success of Lerdahl’s pitch space model can be verified by a comparison of example 3 with the pitch space represented in example 4.

---

14. The evidence for adjudicating the validity of the model is, on the one hand, external: as is well known, modulations were for a period of nearly a century limited to these key areas and these would remain the most common modulatory destinations within tonal music for a century thereafter. More significant, however, is internal evidence: these key areas are heard by listeners as being subjectively ‘closer’ to the tonic than other more distantly related keys. Much of this data is introspectively obvious to anyone familiar with the tonal idiom. The experimental work of Krumhansl and others confirms what most experienced listeners take for granted along these lines. We will return later to a discussion of the important distinction between the external and internal empirical domains.
Rather than being constructed from horizontal third and vertical fifth axes, both axes map out key areas immediately adjacent along the chromatic scale, with the vertical axis alternating, as does Lerdahl’s minor third horizontal axis, major and minor regions.

There is no *a priori* reason why the arrangement of tonal regions in example 4 should not describe the relevant cognitive facts; indeed, there is at least some basis for believing that a fully chromatic space, as example 4 might be called, is not altogether unreasonable as a model of how we orient ourselves in relation to tonal regions. For example, musical instruments tend to situate chromatic pitches adjacent to each other on fingerboards or keyboards. So does conventional musical notation which represents half steps as occurring on vertically adjacent staff locations (with the addition of the appropriate accidentals).

But while there is no obvious *a priori* reason why the space in example 4 shouldn’t succeed in characterizing perceptual reality, the *a posteriori* fact of the matter is that it does not: we simply don’t happen to hear music in that way the model indicates. To return to the previous discussion, the closely related tonic and the dominant key areas, (B and A) are located far away from the tonic E major on the space, requiring a traversal of six vertical or horizontal steps to the left or right respectively. In contrast, regions which are experienced as distant to the tonic, F and Eb major, are represented as immediately adjacent on both axes. Of the closely related keys, only the parallel minor i is represented as close to the tonic, and this is only diagonally adjacent. As for the Chopin major third modulations indicated on Lerdahl’s path these appear as, on the one hand, closer in space than the dominant and subdominant regions I and IV and also more distant than the immediately adjacent chromatic regions 9 (cf. example 3). Thus, this model of our competence clearly fails — not only due to its own empirical shortcomings, but, more significantly, in comparison to “other possibilities within the space”, namely Lerdahl’s pitch space in example 2 which succeeds in describing the data more satisfactorily.

Now it may be the case that the broad class of spatial representations of which Lerdahl’s model constitutes one possible instantiation will turn out to be, for some reason, incapable of accurately representing certain important aspects of
our internalized understanding of tonal hierarchies. Indeed, other organizations have been proposed, most notably the neo-Riemanian model which has become probably the dominant analytic framework for the description of late Romantic repertoire within music theory proper.\footnote{See, for example, Cohn (1999).}

This is not the place to evaluate these competing models except to remark that which representation is chosen is entirely an \textit{a posteriori} empirical question. And in so far as a representation succeeds in explaining the structure which listeners actually assign to it, it is reasonable to assume that, according to Jackendoff’s description, the model successfully “designates what cognitive scientists believe the mind has in it” — a symbolic description which is homologous to the mental structures according to which (the musical experience) is ultimately instantiated.

4. Conclusion

The discussion above allows us to clarify two aspects of the cited remarks by Jackendoff. First, it should be clear that whatever form these representations ultimately must take they are, as Jackendoff insists, by their nature, “not \textit{about} anything”. That is, they do not refer to “things” in the external world independent of our perceptual relation to it but rather designate salient aspects of the construction of particular things — namely pieces of music. As such, the representation is not of an external musical object, but of the internal cognitive structure which the external object gives rise to in the mind of the listener.

It follows that, while indeed “not \textit{about} anything” a mental representation is very much a \textit{thing} in any reasonable sense of the term. Specifically, it is a “natural object”, in precisely the sense discussed in Chomsky’s important essay (Chomsky 2000a). Mental representations are natural objects of a particular sort, that is, they are theoretical constructs, like gravitational fields, atomic particles, or the null subject parameter (in linguistics), whose existence, while not directly apparent to our senses, has been shown to be necessary in order to account for a range of observable empirical facts. The “occult” nature of these objects has, as Chomsky recounts, been repeatedly brought into question over the centuries. By this point in the history of science their existential status is taken as entirely unproblematic and barely discussed. Quarks, gravity, and twenty seven dimensional hyperspace (if it exists) are for the purposes of science, in the words of Steven Weinberg, “real in pretty much the same sense (whatever that is) as the rocks in the fields” (Weinberg 1994). If the mental representations posited by Lerdahl withstand scientific scrutiny they must be accepted as just as unproblematically real.

With this perspective in mind, a second claim made by Jackendoff becomes eminently reasonable: insofar as a musical representation is a natural object (in the Chomskyan sense) “it hardly makes sense to say that the representations one
constructs in response to hearing the Eroica Symphony are true or false.” For while propositions such as “John knows the Eroica Symphony” may be true or false, to say that any natural object is true or false, is a semantic absurdity. One doesn’t say “a chair is true”, “a noble gas is false” or “a mountain range is true”: one says that these objects exist (or not) depending on the evidence provided for them. The same applies to the mental structures which a listener constructs in response to hearing a piece of music. It is not a question of the truth or falsity of the propositional content of these representations, but rather of the existence (positive or negative) of the things themselves. By their nature, they are mute — incapable of expressing anything about the world aside from the fact of their own existence.

To conclude, I will suggest that the confusion surrounding the status of mental representations above is part of a larger picture. The tendency to ascribe content to mental representations of cognitive objects which only grudgingly and unnaturally accept such ascriptions is one characteristic of what Jackendoff describes as the “philosophical” relationship to how our mind construes the world around us and the things in it.

According to Jackendoff, the philosophical approach is concerned with answering the following question:

What is the relationship of the mind to the world, such that we can have knowledge of reality, such that we can have beliefs and desires about things in the world, and such that our sentences can be true and false?

In contrast, the psychological approach, while superficially related, asks a fundamentally distinct question:

How does the brain function as a physical device such that the world seems to us the way that it does, and such that we can behave effectively in the world.

Jackendoff argues that the philosophical approach “in case after case … leads to uncomfortable metaphysical problems (most but not all of them well known) whereas the ‘psychological’ approach permits — at least in principle — a revealing account of the phenomena.”

While I will not develop this aspect of Jackendoff’s discussion, as we begin to discover more about music and our cognitive relationship to it and how we as a species construct our knowledge of the world around us in a variety of experiential domains, it would appear that Jackendoff’s assessment is becoming increasingly uncontroversial.
Bibliographical References


