

Linguistics, cognitive science, and all that jazz¹

PETER W. CULICOVER

Abstract

The question of whether generative grammar offers insights into the mind turns on whether and how a generative grammar is an account of what is in the mind. A potentially useful perspective on this question can be achieved by looking at a cognitive phenomenon that is similar in many respects to language but crucially different, namely jazz. Jazz performance is apparently rule-governed and improvisational, like language. It is useful to take jazz as the exemplar of a complex cognitive task and to think of language as different from jazz in several critical respects that may account for some of its special design features, as well as the fact that it is acquired naturally and without explicit instruction. One difference that may have considerable explanatory force is that language is used to encode and communicate Conceptual Structure, while in the case of jazz (and music in general), what is communicated is the form itself.

1. Questions of language and cognitive science

Cognitive science rose to prominence in the 1970s. At that time, it appeared that it might be possible to find important parallels between the structure of language and the structure of other mental faculties. Language is in the mind, linguistics is the scientific study of language, and cognitive science concerns the architecture of the mind; consequently, it was reasonable to think that what

1. I am very grateful to Bruce Hartung, Ray Jackendoff and two anonymous reviewers for their tolerance of my attempt to open up what is for me, at least, new territory, as well as their challenging and helpful comments on an earlier version of this article. These have led to substantial corrections, reformulations and clarifications, although I may still not have satisfied all of their concerns. Thanks also to the editor for a number of corrections and improvements. The responsibility for any remaining errors and lack of clarity is of course mine alone.

linguistics tells us about the architecture of the mind might turn out to be relevant for other mental faculties. The modularity thesis (Fodor 1983) held out the promise that there might be interesting architectural similarities across a spectrum of cognitive modules, of which language is the most well-studied and therefore to be taken as a prime exemplar. In fact, some suggestive parallels were offered in the domain of vision (Marr 1982)² and music (Jackendoff and Lerdahl 1980).

Given this early promise, it is of some interest to note that some thirty years later, we are considering questions such as those posed by the editor of this volume, such as:

- Is linguistics (generative grammar, or other approaches) a field which truly offers insights into how the mind and brain functions?
- If so, what types of information does it offer that other areas within cognitive science may benefit from?
- Do theories of grammar serve as models for the study of other cognitive systems?
- Specifically, do theories of acquisition of grammar find support in developmental trajectories in other cognitive domains?
- Does language seem to operate in the same way as other cognitive functions do, specifically with regard to memory and learning?

One might be inclined to conclude from the fact that these questions are still being asked at this very general level that linguistics does *not* really offer any insights into how the mind functions, either with respect to language *per se*, or more generally. Such a conclusion has in fact been drawn by many cognitive scientists over the years.³ The response by some linguists who care about the place of linguistics in cognitive science has basically been one of regret that linguists do not do a better job of communicating what the field has discovered about the way the mind works (e.g., Anderson 1989; Jackendoff 1988).

It is no doubt true that many non-linguists find it difficult to understand how a linguistic theory or a particular linguistic analysis tells us something about the mind. A primary reason for this, in my view, is that the role of linguistic theory in cognitive science has not been properly characterized, either within linguistics or outside of linguistics. The fundamental problem is the postulate that a theory of competence must play a central role in a theory of performance. A key formulation of this idea is that of Chomsky (1965): “In general, it seems that the study of performance models incorporating generative grammars may

2. For example, Marr argued for distinct levels of visual representation reminiscent of the levels of representation such as phonology, syntax and semantics (Lerdahl and Jackendoff 1983: 307).

3. For a recent manifestation, and references to others, see Edelman and Christiansen 2003 and Phillips and Lasnik 2003 for a reply.

be a fruitful study; furthermore, it is difficult to imagine any other basis on which a theory of performance might develop.”

One very literal model of how this works is that somewhere in the mind/brain there is a generative grammar that accounts for a person’s knowledge of language, and somewhere else there is a computational device that draws upon this knowledge in order to carry out comprehension and production. I will refer to this device as the performance mechanism.⁴ The particular generative grammar is arrived at in the course of language acquisition through the interaction of the Language Faculty that embodies some theory of competence (UG) and linguistic experience. The performance mechanism computes the representations that are licensed by the grammar,⁵ perhaps by referring directly to the various rules and principles that form the grammar. On this view we would expect to find evidence in performance for particular properties of the generative grammar, which in turn would constitute evidence for a particular theory of competence.

If a theory of competence is incorporated in a theory of performance, then the obvious question to ask is, “How is *this* particular theory of competence incorporated in a theory of performance that actually explains how people learn, produce and comprehend language?” In the absence of good answers to this question, it is natural to be skeptical about the theory of competence. This skepticism could take the form of doubting that a given theory of linguistic competence is relevant to performance, or more generally, that any theory of linguistic competence is relevant to performance.

Competence and performance are in fact intimately related (contrary to the beliefs of many), but I suggest that the traditional view has the relationship backwards. A theory of performance forms the basis for a theory of competence. On this view, what is acquired in the course of language acquisition is performance itself. Knowledge is embodied in the performance. It develops as the range of performance grows and generalizes. A generative grammar is an idealized description of this knowledge. The theory of competence constrains possible descriptions; it is a theory of the knowledge that can be embodied in the performance.

There are, of course, many ways in which we might understand how a theory of competence might be incorporated in a theory of performance, some of which do not entail that we would be able to find evidence for the theory of competence through the examination of performance. For example, there could

4. I am using the term “performance” here in the sense of “processing.” It is logically possible and in fact plausible that there are two performance mechanisms, one for production and one for comprehension. I put this possibility aside in order to keep the exposition as simple as possible.

5. I will use the term “grammar” here and subsequently to mean “generative grammar.”

be a maturational process by which a performance mechanism is constructed on the basis of a generative grammar, after which the generative grammar simply withers away.

Or it could be that a generative grammar *per se* is not constructed in the mind/brain under the guidance of UG; only a performance mechanism is constructed that gives the appearance that there is a generative grammar in the mind/brain. This idea is related to the suggestion by Jackendoff (2002: 426) that UG “guides the course of the child’s generalizations over the evidence.” These generalizations may be embodied in the performance mechanism, and the linguist can seek to describe them in formal terms, but there is no need to assume that they are represented in the mind/brain in a form that resembles the description.

I think that this is basically the right way to look at the issue. It leaves open the possibility that some aspects of the knowledge that is embodied in the performance mechanism can be accounted in terms of some account of UG, while others arise in other ways. It is these other ways that I will be primarily concerned with here, through the comparison with jazz.

I start from the usual position that language is in the mind, that cognitive science is the study of the architecture of the mind, and that linguistics must therefore be part of cognitive science. Taking this view seriously, I compare language with perhaps its closest parallel, that of jazz, with the goal of understanding what is special about each of them and what they have in common. Jazz is in the mind, too, and thus must fall under cognitive science. This simple observation leads to the hypothesis that jazz might in some ways be a good model for language. A consideration of how jazz is learned, how it is represented in the mind, and how improvised jazz performance is possible might suggest answers for similar questions about language. If so, we might expect to find that some of the differences between language and jazz can be accounted for in terms of the basis for their acquisition and their different communicative functions. I will suggest some reasons to think that this line of reasoning may be on the right track.

2. Jazz as a “language”

Jazz musicians frequently speak about jazz as a “language” and draw parallels between learning to play jazz and learning a language. For example, there is a video biography of Toshiko Akiyoshi entitled *Jazz Is My Native Language*, and there are books entitled *The Jazz Idiom* (Coker 1975) and *The Jazz Language* (Haerle 1980). Musicians liken the acquisition of jazz competence to the acquisition of a language. Here is a typical example from a *Jazz Improv* interview with the trumpeter Valery Ponomarev (Nemeyer 2004).

VP: Rhythmic, harmonic, melodic, form-wise. It's all language, and, to teach, or to learn, to speak jazz is identical to learning to speak any language. Just like a baby starts with first sounds ... on and on, right? And then he goes on and on, until all those speech particles become words, become a string of words, become sentences. And then the person's just talking. Right now, I'm not talking ... thinking of grammar rules ...

JJ: And you'll transpose that to all keys and you know, the players will plug in their little licks because it works. It's ok, but they're not really saying anything until you get to the point like you and I were just extemporaneously talking here – in the same way that when we improvise, or play, together we would do the same thing musically.

VP: Exactly. At one point, it all goes to subconscious. And you are playing music.

In the same vein, Berliner (1994) begins Chapter 4 of his extensive study of jazz performance and culture entitled “Getting Your Vocabulary Straight,” as follows:

Just as children learn to speak their native language by imitating older competent speakers, so young musicians learn to speak jazz by imitating seasoned improvisers. In part this involves acquiring a complex vocabulary of conventional phrases and phrase components, which improvisers draw upon in formulating the melody of a jazz solo. (1994: 95)

Jazz musicians often talk about “saying something” when they improvise (Monson 1996). Some draw the analogy with spoken language even more closely, using terms such as ‘vocabulary’ and ‘grammar’ and comparing improvisation with speaking.

Marvin Stamm:⁶

In truth, improvisation is a musical language with which Jazz musicians converse with each other. They learn just like everyone else who masters their native language; not from books, but by ear! ... the young musician develops a vocabulary and learns the intricacies of the language, the phrasing, articulations, rhythms and harmonic approaches in order to clearly express himself. ... The more expansive his vocabulary, the more fluent he becomes in the language ... Over a period of time, the neophyte player develops into a full-blown Jazz linguist, and, if perceptive, realizes this is a life-long task in which he will constantly participate in to become more fluent, so he can enhance his musical vocabulary, find his own highly-individualized voice, and establish a niche for himself with his music ... I play extemporaneously, just as we talk extemporaneously. Do we think about speaking? Or, isn't our speaking dictated by the

6. <http://www.52ndstreet.com/kenton/stamm.htm>.

situation at hand and based on our need to respond to that situation by expressing our own ideas and thoughts, which then creates a dialogue with those who elicited that response. Jazz improvisation, like speaking, is just that.

Garrett Leaver:⁷

Jazz Vocabulary : whole tone scale, bebop scale, diminished scale, altered chord tones, pentatonic scales, hypo dorian, nice sounds – but now it is time to apply them to various chord progressions to make them not sound like “scales” running up and down your instrument but like a *melody* that expresses how you feel. In my last few sentences I didn’t suddenly think “oh shoot . . . I haven’t used an adverb for 10 minutes, I better get one in quick(ly)!” If I needed an adverb to express what I feel, I use it as part of my natural language.

Course description, University of Alabama:

Jazz Piano/Jazz Theory I – MUS 228⁸
 Moody Music Building Room 249 – 1 Credit Hour
 Prerequisite: MUS 217 or permission of the instructor

Purpose/Objective

The objective of Jazz Piano/Jazz Theory I is to introduce the student to the fundamental grammar, vocabulary and structure of the jazz language through the study of its nomenclature, chords, scales and their application on the piano.

Of course these are not theoretical statements nor should we expect them to be, given that jazz musicians are not theoretical linguists. It would be a mistake to try to understand them as claims that the structure of jazz is essentially that of a natural language. But they do capture the intuition that jazz musicians have quite generally, based on their experience in acquiring and performing the idiom, that jazz has some key properties that make it resemble a language.

3. How jazz is like language

The experiences of jazz musicians suggest that there may be something important for the cognitive scientist to learn by studying what constitutes the capacity to perform jazz, and finding points of comparison with language. One main reason to do this is to use the insights about how non-linguistic tasks are acquired, represented, and performed as a window on how language might be acquired, represented and performed. Another is to be able to understand what is special about language, and why, through an analysis of what is found in language but not in other domains.

7. <http://www2.magma.com/~garrett/>.

8. <http://bama.ua.edu/~jnoffsin/mus228/>

Let us try to enumerate some of the closest points of comparison between jazz and language.

- *Jazz is improvised.* Johnson-Laird (2002) puts it as follows: “If you are not an improvising musician, then the best analogy to improvisation is your spontaneous speech. If you ask yourself how you are able to speak a sequence of English sentences that make sense, then you will find that you are consciously aware of only the tip of the process.”
- *Jazz appears to be “rule-governed”.* In a particular jazz style, e.g., bebop, certain sequences of elements are judged by competent practitioners to be ill-formed, while others are judged to be fully acceptable. It is possible to state fairly precisely what constitutes a well-formed sequence, and experienced musicians share these intuitions.
- *Jazz is processed in real time.* The player is able to produce well-formed improvisations on the fly, without rehearsal (although of course in both cases there must have been considerable prior practice in order to make this possible, a point to which I return.)
- *Jazz is creative.* In the course of jazz performance, a musician will typically produce sequences that satisfy the constraints and have never been produced before (at least not by that musician).
- *Jazz is acquired through experience.* Like language, jazz is not taught explicitly, although there are problems that a jazz musician may face that can be solved through analysis and instruction.
- *Jazz is used for communication.* Jazz does not communicate propositions, but it does communicate aesthetic characteristics and emotions.

There are, of course, many important ways in which jazz is *not* like language; these will become more pertinent after we have given some additional consideration to some of the parallels. But there is one that is important to note at the outset. In the case of language, meaning is communicated through apprehension of the form of an expression. When we process a sentence with a particular structure, we compute the structure as a means of determining the meaning. In the case of jazz,⁹ what is communicated is the form itself.¹⁰ When we process a piece of music, we compute just the structure. There is no content associated with the parts, and the whole has no meaning. Communication consists of successful recognition of the structure on the basis of the musical elements.

Here is a simple example to illustrate this point. A typical jazz performance includes a set of improvisations of a piece of written music. This music has a structure and a melody. The structure is a sequence of chords and the melody

9. I am referring here to the more traditional styles of jazz in which musical structure has not been explicitly abandoned.

10. In both cases, emotion can be conveyed through the manner of expression.

is a sequence of notes. Each chord consists of specific notes and has a duration, and the subsequence of notes that occurs during the duration of a given chord has a particular relationship to the notes that belong to the chord.

Consider Figure 1; from the bottom to the top the notes C, E, G and B \flat are the root, 3rd, 5th and dominant 7th respectively of the C dominant 7th chord. The first measure shows the four notes of the chord. The second measure shows a simple melody consisting of exactly these four notes in a particular sequence. There are four beats to this measure and each note in this measure occurs on a beat.



Figure 1. C7 chord

The rough linguistic analogy is that the chord corresponds to a phrase of a particular category, the sequence of chords corresponds to the phrase structure of a sentence, and the melody corresponds to the sentence itself.

There are a number of ways to realize a particular phrasal category in a particular position in a sentence, and there are a number of ways to realize a particular chord in a particular position in a musical piece. In an improvisation over a structured piece of music, such as a popular song, or a song from a Broadway musical, or a traditional folksong, or a newly composed piece, the sequence of chords is fixed and aligned with the temporal structure of the piece. What is improvised is the melody, and what is important is that the melody at each point in the song be consistent with the chord that is in effect at that point. Crucially, not every note in the melody has to belong to the chord, but there are rules. One fundamental rule (which of course may be broken for aesthetic purposes) is that a beat, particularly a strong beat (the first and the third beat of a measure) should be a chord tone. So if we play two notes per beat, we want to align the sequence so that the note that is aligned with the beat, typically the first note, is a chord tone. In the first measure of Figure 2 this rule is followed; in the second measure, it is not. (The tones on the beat are circled.)



Figure 2. Chord tones on strong beats (measure 1) and not on strong beats (measure 2)

The first melody is a much better realization of this chord sequence than the second.

It is possible to put other tones on the beats, a practice that adds color and melodic interest to the melody that is created. Some of these non-chord tones are called “color tones”. In the key of C, the color tones are D (the 9th), F (the 11th), and A (the 6th or 13th); these are the tones that appear on the first three beats of measure 2 in Figure 2. More color can be added in certain contexts through the use of other tones, particularly the sharp 11th, the flatted 9th, the sharp 9th, and the flatted 13th. Typically at least one basic chord tone will fall on a beat in a measure, however.

It is relatively straightforward (although not trivial) to construct melodies on paper such as measure 1 of Figure 2 that satisfy the rules. It is quite another thing to create melodies in real time, at high speed, as the piece moves through a sequence of chords, such that the melody is aesthetically pleasing, the transitions from one chord to another are perceived as smooth and natural, and the rules are satisfied.

Comparable improvisation is of course what is done during the course of producing language. The jazz musician’s intuition is that the way that it is done in jazz is that the musician learns a “vocabulary” of ways to realize certain forms that satisfy the various constraints. Groupings of notes, arranged in a particular way, are analogous to the sounds of a language that constitute a word or phrase when arranged in a certain way. The elements of the jazz vocabulary are practiced until they become automatic. In the course of improvisation they are inserted into the sequence to express the player’s communicative intention. An illustration is given in Figure 3, which is a short improvised solo that illustrates the linking together of phrases as the chords change. The beats are circled. Only the Fs in the first two measures are chord tones; G, C and B \flat in the third measure are chord tones.



Figure 3. Solo #1

Note the smooth local transition from E (the 9th of D minor (or Dm)) in measure 1 to F (the dominant 7th of G) in measure 2, and the smooth local transition from E–E \flat (the 6th of G) to D (the 9th of C) in measure 3.

For the sake of comparison, Figure 4 shows an entirely different solo over the same chords. Non-chord tones are used in the Dm7 chord. Again, note the smooth local transitions.



Figure 4. Solo #2

It is particularly important to consider how the vocabulary of jazz may be represented internally. As noted, a crucial requirement is that chord tones fall on the strong beats. This can be accomplished in a number of ways, because at any given point in a harmonic sequence, a given tone is a member of many chords, including the chord that is written in the music. Consequently, it is possible to play a pattern based on a chord other than what is written in the music and produce a melody that is consistent with this basic requirement. Since the substitute chord is not identical to the written chord, some of the tones will be color tones of the written chord even as they are chord tones of the substitute chord.

Here is an illustration. Suppose that we want to play a melody over the chord $\text{Em}7\flat 5$, shown in Figure 5. The notes are E (root), G (minor 3rd), $\text{B}\flat$ (flatted 5th) and D (dominant 7th).

Figure 5. $\text{Em}7\flat 5$

The chord $\text{GMaj-m}7$ consists of G (root), $\text{B}\flat$ (minor 3rd), D (5th), and $\text{F}\sharp$. Note the overlap; the $\text{F}\sharp$ of the $\text{GMaj-m}7$ chord is the 9th of Em, a color tone. Figure 6 shows a melody over $\text{Em}7\flat 5$ that utilizes the tones of the $\text{GMaj-m}7$ chord.

Figure 6. Solo over $\text{Em}7\flat 5$

These examples are intended to illustrate the fact that it is possible to learn, internalize and make automatic a repertory of gestures that will produce well-formed sequences. The correctly chosen variant of a certain gesture, for exam-

ple, the GMaj–m7 over Em7b5, is, in a sense, a way of realizing or “saying something” in that chord.

4. Where is the grammar of jazz?

As in the case of language, we can ask, where does competence reside? Is there a generative grammar of jazz in the mind/brain that is somehow “incorporated” in the jazz performance mechanism? Or is the description of what the jazz musician is doing in an improvisation, along the lines of what is sketched out in Section 3, a description of the knowledge that is embodied in the musician’s real-time capacity to improvise? I would like to suggest that it is the latter. As far as the mind/brain is concerned, the grammar of jazz resides in the performance device. This performance device contains a set of gestures and ways of linking gestures together to realize certain chords and sequences of chords. The rule that says that a strong beat should contain a chord tone is not appealed to in the course of improvisation, but is embodied in the set of gestures that will produce the consequence that a strong beat contains a chord tone.

In order to improvise a melody, the musician must have internalized a set of *gestures* over a range of chords, and must be able to produce the appropriate sequence of gestures for each chord in the sequence of the piece. A gesture may be a single interval made up from a sequence of two notes, or a linked set of intervals in sequence. In a sense, the sequence of chords is similar to the phrase structure of a language. A complex expression of a specific type, e.g., a sequence built over a Dm7 chord, precedes an expression of another type, e.g., a sequence built over a G7 chord, as in Figure 4. There is an infinite number of well-formed sequences that can be built over any chord, given possible variations in melody, timing, rhythm and so on.

It is therefore natural that the linguist may see in the structure of jazz something that resembles a grammar. Given that many of the rules do not appear to be hard and fast, but simply preferences that may be overridden in some contexts by other preferences (e.g., the preference for “color” may override the preference for a chord tone on a strong beat), the linguist might even propose that the correct grammar of jazz is an optimality theoretic one. Johnson-Laird draws the parallel as follows:

In modern jazz, each improvisation is novel in the same way that most utterances in natural language are novel – the speaker has never uttered them before, and often neither has anyone else. As far as one can tell, improvisation is not deterministic, that is, it does not unwind like clockwork with only one choice for the musician at each point in the performance. There are harmonic and rhythmic constraints on improvisations. Much of the hard work in learning to improvise consists in acquiring a tacit mastery of these constraints. The existing elements

are the fundamental constituents of tonal music: pitches, timbres, chords, intensities, and durations. (2002: 420)

Just as we might characterize the musician's knowledge in terms of a grammatical description of some type, we might seek to account for the development of the jazz musician's ability to improvise in terms of the development of such a grammar, perhaps starting from some primitive task-specific basis consisting minimally of the ability to distinguish musical tones and rhythms and the ability to produce them using the appropriate articulators.

As in the case of natural language, it is natural to expect that it would be possible to state explicitly what the rules of the grammar are in the case of jazz. We would expect to be able to use these rules to produce an infinite number of improvisations over a set of chords. In fact there is at least one computer program that does just this, called *Band-in-a-Box*[®]. It not only creates solos, but it can also create chord sequences and melodies over those chord sequences. It can do this in a large number of styles, using the clichés and mannerisms (that is, the vocabulary) of well-known musicians. For each chord, *Band-in-a-Box*[®] produces a phrase that makes a natural transition from the preceding phrase. The result is a smooth sequence of phrases that satisfies the constraints. Figure 7 shows the melody of the beginning of *Happy Birthday To You*, and Figure 8 shows a solo generated over the same chords.

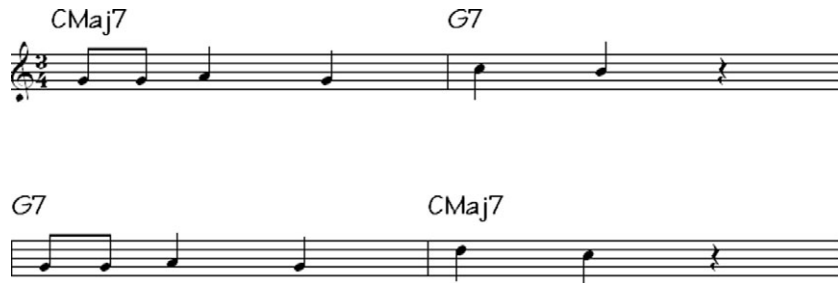


Figure 7. First four measures of *Happy Birthday To You*

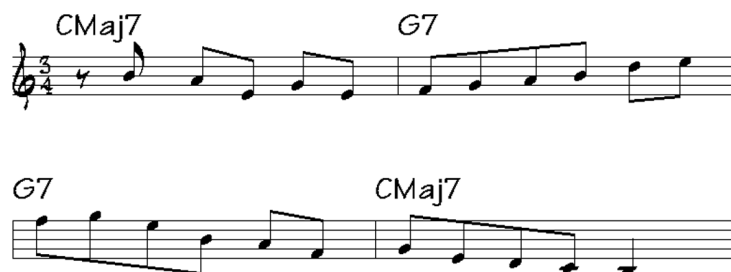


Figure 8. Solo over chords of first four measures of *Happy Birthday To You*

In the solo, the E at the end of the first measure, which is the 3rd of CMaj7, leads to F, which is the dominant 7th of G7. Similarly at the end of the second measure. The sequence A–F in the third measure leads back to the G of CMaj7, which is the 5th, by circling around it.

So the parallels between language and jazz seem fairly clear.¹¹ But it does not seem immediately plausible to posit a UG specific to jazz that guides the formation of a grammar or of a performance mechanism that embodies the knowledge that such a grammar describes. I elaborate this point in the next section.

5. The grammar of improvisation

It is clear that there is substantial creativity in the performance of jazz, and that what is created is subject to constraints. It is also possible to describe to some level of accuracy what these constraints are and to give an account of what constitutes acceptable and unacceptable jazz performance. *Band-in-a-Box*® successfully produces well-formed improvisation, so one might be tempted to suppose that the musician has in his or her head something similar that carries out the same function.

At the same time, it is important to recognize how something like *Band-in-a-Box*® comes about and what might be different about the musician's capacity for improvisation. In order to construct a program like *Band-in-a-Box*®, it is necessary to examine what musicians do, describe it in general terms (that is, figure out what the rules are), and then implement the rules. There is clearly some relationship between the description and the program, in the sense that the program produces the sequences permitted by the rules. At the very least,

11. For extensive discussion of the formal parallels between language and music, see Lerdahl and Jackendoff (1983).

the rules are referenced in the course of writing the computer program that produces the solo.

Similarly, the knowledge that the musician has is acquired either through induction on the basis of exposure to exemplars, or through explicit instruction. It is possible to state in terms of a list of rules what the musician knows, at least approximately. It is logically possible, but not particularly plausible, that the theorist's description corresponds to something that is actively engaged in the performance of jazz. Aspects of the theorist's description may be invoked in the course of teaching and learning, as a way of defining properties of the patterns that the musician needs to practice in order to become adept as an improviser. But such descriptive statements are not the same as accounts of how the improviser performs in real time, which does not appear in general to involve explicit reference to these rules.

Rather, it appears that the theorist's description is an account of the general properties of the musical relationships that are produced by the performance mechanism for jazz. The theorist observes the generalization that it is necessary or preferable that chord tones fall on the strong beats; what is in the musician's head reliably produces this effect and we can truly say that this is something that the musician "knows". But crucially, what the musician knows concerns the nature of correct *performance*. The description of the properties of this performance, which we might call the "theory of competence," is secondary. In this sense, we may reasonably say in this case that *performance is the basis for a theory of competence*. The converse, that the theory of competence is incorporated in a theory of performance, has little explanatory role in this domain. There is no evidence to suggest that a theory of competence along the lines envisioned for language precedes performance either logically, developmentally, or in real-time performance.

To put it another way, there does not appear to be any basis for assuming that there is a "Jazz Faculty" that is part of our biological endowment for music, which defines a set (possibly infinite) of "jazz grammars," of which the actually occurring ones (that is, those in the heads of actual jazz musicians) are drawn. Nor does there appear to be any basis for suggesting that the acquisition of jazz involves the fixing of a set of hardwired parameters, or any other process that involves plotting a trajectory through a (biologically) predetermined hypothesis space on the basis of experience.¹² Rather, what appears to be going on in the acquisition of jazz is what many jazz musicians have suggested: "the young

12. Of course, there is very likely a universal capacity for the general apprehension of music (Lerdahl and Jackendoff 1983), and some particular cognitive capacity that allows for the ability to perform jazz. Some people have more of this capacity than others. (I certainly don't have anything close to as much as I would like.) I do not believe that this constitutes a basis for arguing that there is an innate capacity for jazz *per se*, however.

musician develops a vocabulary and learns the intricacies of the language, the phrasing, articulations, rhythms and harmonic approaches in order to clearly express himself" (Stamm) based on experience and practice. Jazz competence is constructed out of jazz performance.¹³

It appears to be possible, in principle at least, to describe the knowledge that users have in formal terms, in the case of both jazz and language. But there appears to be no strong argument, in the case of jazz, that the architecture of the knowledge resembles in any principled way the formal properties of the description. For example, there is no evidence that the knowledge that the beats should be chord tones is directly drawn upon in the course of performance, although it may be during instruction and practice. Rather, the jazz musician executes various learned but not invariable gestures in performance that have the consequence that certain beats will be chord tones, or color tones, with certain aesthetic effects.

The study of jazz suggests that the capacity to engage in complex performance can be acquired through experience and instruction, it can be internalized, and it can be used to produce novel instances in real time. Given that this is the case, it is reasonable to suppose that the same might be true in the case of language, at least to some extent. It might be that what is acquired in the case of language acquisition is not a generative grammar in the theoretical sense (see, for example, Culicover 1997). Rather, it is a complex performance mechanism that under various idealized descriptions characterizes an infinite set of sequences that are well-formed (and perhaps meaningful), or that more or less optimally satisfy some set of constraints. In this sense, the theory of performance in language forms the basis for a theory of competence. As Nowak and I have argued elsewhere (Culicover and Nowak 2003), on this view a grammar is not a description of the device in the head that is responsible for language. It is a description to a certain level of abstractness of the behavior of a complex dynamical system that is in the head whose function is to compute correspondences between form and meaning.

But if we take this step, which seems reasonable, we are immediately faced with the inescapable fact that jazz, unlike language, is not something that everyone learns natively. The assumption that there is a Language Faculty in the mind (but not a Jazz Faculty), immediately explains this difference (if you call a stipulation an explanation). If we abandon the strongest nativist position about language, we still have to account for the obvious differences between the two tasks, and particularly the fact that every person learns a language, but not every person learns jazz.

13. See Quartz and Sejnowski (1997) for a characterization of this 'constructivist' approach to learning. Berliner (1994) contains the best published description of the process through which jazz musicians acquire the jazz idiom.

6. Where does this lead?

Suppose that we take jazz improvisation as the exemplar, and try to see language as a similar phenomenon in which very different types of relations are computed. From this perspective, we would want to understand not only the parallels between the tasks, but the differences, and we would want to see if the differences between the tasks might help to explain why they clearly have a different status with respect to acquisition. Here are the areas of some particularly salient differences between language and jazz improvisation.

- *The nature of the knowledge and of what is communicated.* Expressions in a natural language correspond to Conceptual Structure; expressions in jazz do not.
- *The types of relations represented.* Natural language encodes scope and other non-adjacent and non-local relations. Jazz does not, nor does music more generally.
- *The basis for acquisition.* Language is acquired on the basis of massively redundant information; jazz is not nearly as ubiquitous for the developing child.

I address each in turn briefly in the following sections.

6.1. Knowledge

Crucially, knowledge of language consists of knowledge of form-meaning correspondences.¹⁴ Knowledge of jazz consists of knowledge of possible forms; there is no meaning. In the domain of ordinary language, form is relevant as the means by which meaning is conveyed.¹⁵ The objective is to express and communicate literal meaning (semantics), non-literal meaning (presupposition and implicature) and perspectives on the ongoing discourse (topic and focus). The biological imperative to learn language may be cast in terms of a biological imperative to communicate desires, intentions, and ideas.¹⁶

In jazz, on the other hand, there is no meaning beyond the form.¹⁷ What is communicated *is* the form. The listener appreciates the form: the beauty of

14. See Jackendoff (2003) and Culicover and Jackendoff (2005) for a particular view about what these correspondences consist of.

15. Poetic and literary language may be produced with the intention of communicating ideas and at the same time with the intention that the form of expression be recognized as form.

16. See Tomasello (2003). Whether we need to posit a specific imperative to map sounds to meanings, or whether this imperative follows from a more general cognitive imperative to discover correlations in the world wherever they occur, is an open question that I cannot begin to explore here.

17. See Kivy (1993) for a related view of music in general.

the melody, the originality, the humor, the energy, the technical achievement. Of paramount importance is the player's ability to produce a melody over a fixed or flexible sequence of chords that satisfies the constraints of the form as well as aesthetic criteria.¹⁸ Since jazz, and music more generally, lacks the communicative function of language, it should not be surprising that it is not learned by all human beings upon exposure.

At the same time, there is always a question of what knowledge is acquired and what knowledge must be assumed to be innate. It does seem to be the case that there are formal properties of language that learners know that they cannot discover simply from examining the linguistic evidence presented to them. Among the properties suggested by Culicover and Nowak (2003) are that phrases are projections of heads (the irreducible core of X' theory) and that there can be discontinuous dependencies between a dislocated phrase and its governor. Culicover and Jackendoff (2005) argue that a representational tier that contains the grammatical relations Subject and Object is an ineliminable architectural component of the mapping between sound and meaning. But as always, the challenge is to try to find a way that such knowledge can be acquired strictly on the basis of experience, in spite of its apparent unlearnability, in order to ultimately sustain the claim that it must be part of the innate linguistic capacity.

As noted in Section 1, even if we conclude that a particular aspect of linguistic knowledge is innate, we are not required to abandon the view that what is in the head is a performance mechanism and not a generative grammar *per se*. Rather, there may be some innate mechanism for constructing a performance mechanism on the basis of exposure to linguistic data, consistent with certain constraints.

6.2. *Function*

It is possible to say more than simply that language is used to convey meaning, and jazz is not. Language expresses Conceptual Structure (CS) (Jackendoff 1983, 1990, 1997). It should not be surprising that the syntactic structure of natural languages approximates that of CS.¹⁹ Consider the computational

18. It is often suggested that a central function of music is to communicate emotion. This is a property that music shares with all other forms of expression, including language and dance. It seems plausible that the absence of propositional content may serve to highlight the emotional force of artistic expression. In fact, many musicians feel that performance of vocal music constrains the expressive possibilities because of the restrictions imposed by the text (Bruce Hartung, p.c.).

19. For an extended discussion of this idea, see Culicover (1999). I suggest there that in many respects what has been thought of as UG is really CS.

requirements of mapping between form and meaning, and let us accept what is generally assumed, which is that the computational system seeks economy wherever possible. Intuitively, the form-meaning correspondence is most economical when there is a single form for a single meaning or component of meaning. Moreover, the mapping is maximally simple when the relationships between the parts are maximally transparent; see Culicover and Nowak (2003). In maximally simple cases, recognition of the form immediately triggers the corresponding meaning, and vice versa.

The form that is expressed in jazz is essentially that of melodic and rhythmic variants over a sequence of chords. Johnson-Laird (2002) makes the case that this task can be accomplished locally, without implicating working memory. The rhythm and melody produced from a given point in the sequence can be accounted for in terms of the immediately preceding rhythm and melody. There is no requirement that aspects of the sequence must be put into a memory store so that they can constrain the form at some later point in the sequence. As Johnson-Laird points out, this is different from the fact that it is possible for musicians to intentionally store a melody and repeat it later, or produce variations on it. The claim of local dependence is that there is nothing about the form that requires reference to some earlier part of the sequence.

Contrast this with *wh*-questions in natural language. In English, for example, a *wh*-question is typically formed by positioning the *wh*-phrase in clause-initial position. The fronted *wh*-phrase must be matched by the *absence* of a phrase of the same category later in the sentence where such a phrase would otherwise be required or licensed.

- (1) a. What *are you looking at* ____ ?
b. *What *are you looking at* my shoes?
- (2) a. Why *are you smiling* ____ ?
b. *Why *are you smiling because you're happy*?

In English, an interrogative noun phrase such as *what* completely lacks information about its grammatical function, and hence its semantic function cannot be determined until the “gap” is encountered in the sentence.²⁰ On the other hand, the semantic content of the fronted constituent is constrained by the later verb of which it is an argument – there is selection. So (3a) is fully acceptable but (3b) is problematic.

20. As suggested early, at the very least the *wh*-phrase has to be associated with its governor, that is, the head that determines its grammatical function. There are a number of formal devices through which this linking can be expressed; association with a gap is a particularly transparent one.

- (3) a. *What kind of beer are you drinking* ____ ?
 b. **What kind of book are you drinking* ____ ?

An adverbial like *why* does convey its semantic function directly, but there is still a “gap” because it is necessary to determine the expression over which this semantic function is to be applied.

This dependency between the *wh*-phrase and the gap is a matter of linguistic form, not a matter of style. Such dependencies are characteristic of natural language. It is reasonable to tie the fact that such dependencies exist to the fact that they express directly a function-argument relationship that holds at CS; informally,

- (4) a. WH-X[LOOKING-AT(YOU,X)]
 b. WH-Y[SMILING(YOU, FOR REASON Y)]

The idea is that certain aspects of CS may be represented more or less directly in syntax, which reduces the computational burden for those aspects. Of course, the computational requirements for such dependencies are complex and there are tradeoffs. In English, the scope of the *wh*-operator is marked explicitly, while computation is required to find the appropriate gap. In other languages, such as Japanese, the scope of the operator is marked by a neutral particle, the operator is in situ, and what must be computed is the association of the operator with the particle (characterized in terms of “LF movement” in Principles and Parameters Theory). Additional complexities arise with topicalization, where the topicalized expression expresses a discourse function, not a logical function, but here again, there are expressive correlates to the form.²¹

Such relationships do not exist in music, improvised or otherwise, as far as I know. We might speculate that the computational demands of such relationships would be too great in a strictly formal system, one that does not serve the function of expressing meaning. There is an intuitive basis for this speculation, given that a processor that is translating a form into a meaning must have an intermediate working memory for the meaning as it is constructed. Nothing in principle rules out the existence of such working memory for a purely formal system such as jazz, but nothing about the task requires its existence, either.

6.3. *The basis for acquisition*

Language is acquired on the basis of massive exposure to exemplars (which typically refer to a relatively limited universe of discourse in language directed

21. For a theoretical development of these ideas, see Culicover and Jackendoff (2005).

to children) that taken together provide overwhelming statistical evidence both for constraints on form and for the form-meaning correspondences. Jazz is also acquired on the basis of exposure to exemplars, but only rarely will the experience even approach the density of the linguistic experience. The language learner has ample opportunity to observe not only that there are massive statistical regularities in the forms that he or she is exposed to in the auditory domain (or visual, in the case of sign language), but that there are also massive correlations between these forms and properties of the world, mediated by Conceptual Structure representation. For example, in the case of a set of sentences with the word *cat* spoken in context, not only is there a non-accidental sound sequence [kæt] that can be identified and extracted from a corpus of utterances by a statistical analyzer, but there is also likely to be a high correlation with scenes in the world that contain a cat or the image of a cat. This is the case whether or not the language is being directed to the learner.²²

In contrast, in the case of jazz we have only the form. The situation is analogous to one in which a language learner is given sentences without meanings. It is possible, given certain constraints on the possible hypotheses that can be formed, or taking into account the statistical regularities found in the input, to identify a grammar that will generate the language from which the exemplars are taken and more. If there is a disconnect from meaning, the language acquired will have a decidedly impaired character, such as we find in the case of conduction aphasia or possibly Williams Syndrome. In these cases, the linguistic expressions are well-formed but nonsensical.

7. Conclusions

I have suggested here that not only is linguistics still part of cognitive science, but that we can learn a lot about cognition by thinking about language along with comparable cognitive domains. The one that I chose is jazz, which is not a traditional point of comparison, but proves to have a number of properties that make it an unusually sympathetic accompanist. Perhaps most significant is that language communicates content, while jazz communicates form. The communication of content is arguably of considerably more value to the survival of the species than the communication of form. It should thus not be particularly surprising that children universally acquire a language as a vehicle for accomplishing their goals, while the acquisition of jazz is much more restricted. Taking jazz as an exemplar, we can make the argument that generative

22. In Kaluli society, for example, adults do not speak directly to young children; nevertheless, children acquire the language that is spoken around them as well as children who grow up in cultures where talking directly to children is the norm (Schieffelin 1985).

grammars are useful as descriptions of the behavior of the complex system that is in our heads, but that what is really in our heads is a mechanism for performance that forms the basis for an account of competence. Hence, to answer one of the questions that form the basis for these contributions, “Do theories of grammar serve as models for the study of other cognitive systems?”, I would suggest that the study of other cognitive systems, such as jazz, might well serve as models for theories of grammar.

The Ohio State University

References

- Anderson, Stephen R. (1989). Review of *The Computer and the Mind: An Introduction to Cognitive Science*, by Philip Johnson-Laird. *Language* 65: 800–811.
- Berliner, Eric (1994). *Thinking in Jazz*. Chicago: University of Chicago Press.
- Chomsky, Noam (1965). *Aspects of the Theory of Syntax*. Cambridge, MA: MIT Press.
- Coker, Jerry (1975). *The Jazz Idiom*. Englewood Cliffs, NJ: Prentice Hall.
- Culicover, Peter W. (1997). *Principles and Parameters: An Introduction to the Theory of Government and Binding and Its Extensions*. Oxford: Oxford University Press.
- (1999). *Syntactic Nuts*. Oxford: Oxford University Press.
- Culicover, Peter W., and Ray Jackendoff (2005). *Simpler Syntax*. Oxford: Oxford University Press.
- Culicover, Peter W., and Andrzej Nowak (2003). Markedness, antisymmetry and the complexity of constructions. In *Linguistic Variation Yearbook, Vol (2)*, Pierre Pica and Johan Rooryk, (eds.), 5–30. Amsterdam: John Benjamins.
- Edelman, Shimon, and Morten H. Christiansen (2003). How seriously should we take minimalist syntax? *Trends in Cognitive Sciences* 7 (2): 60–61.
- Fodor, Jerry A. (1983). *The Modularity of Mind*. Cambridge, MA: MIT Press.
- Haerle, Dan (1980). *The Jazz Language: A Theory Text*. Warner Bros Publications, Inc.
- Jackendoff, Ray (1983). *Semantics and Cognition*. Cambridge: MIT Press.
- (1988). Why are they saying these things about us? *Natural Language and Linguistic Theory* 6: 435–442.
- (1990). *Semantic Structures*. Cambridge, MA.: MIT Press.
- (1997). *The Architecture of the Language Faculty*. Cambridge, MA: MIT Press.
- (2003). *Foundations of Language*. Oxford: Oxford University Press.
- Johnson-Laird, Philip (2002). How jazz musicians improvise. *Music Cognition* 19 (3): 415–442.
- Kivy, Peter (1993). *The Fine Art of Repetition*. Cambridge: Cambridge University Press.
- Lerdahl, Fred and Ray Jackendoff (1983). *A Generative Theory of Tonal Music*. Cambridge, MA: MIT Press.
- Marr, David (1982). *Vision*. San Francisco: WH Freeman and Co.
- Monson, Ingrid (1996). *Saying Something Jazz Improvisation and Interaction*. Chicago, IL: The University of Chicago Press.
- Nemeyer, Eric (2004). Interview with Valery Ponomarev. *Jazz Improv* 4.3: 108–114.
- Phillips, Colin, and Howard Lasnik (2003). Linguistics and empirical evidence. *Trends in Cognitive Science* 7 (2): 61–62.
- Quartz, S. R., and T.J. Sejnowski (1997). The neural basis of cognitive development: a constructionist manifesto. *Behavioral and Brain Sciences* 20 (4): 537–596.
- Schieffelin, Bambi B. (1985). The acquisition of Kaluli. In *The Crosslinguistic Study of Language Acquisition, Volume 1: The Data*. Dan Slobin (ed.), 525–594. Hillsdale, N.J.: Lawrence Erlbaum Associates.

Tomasello, Michael (2003). *Constructing a Language: A Usage-Based Theory of Language Acquisition*. Cambridge, MA: Harvard University Press.