MR. DAVID BRAINARD: Good afternoon and welcome. I'm David Brainard, the director of the Institute for Research in Cognitive Science here at Penn. And I'm pleased to welcome you to the 15th Annual Benjamin and Anne Pinkel Endowed Lecture.

Before I introduce the lecture series, I want to mention that at the end of the talk, there'll be a reception up there. And in fact, out there are some kind of Penn previews things, and the staff will be standing fiercely at that door, not letting you out, and directing you up this way. So just save yourself one step by going up the stairs and exit at the end.

I am going to tell you a little bit about the lecture series. It was endowed in honor of Benjamin and Anne Pinkel by their daughter, Sheila, in their memory, to remember them. And each year, we bring in a speaker in the general area of cognitive science, whose interests are matched to those of Benjamin Pinkel, who himself was very interested in matters of consciousness and the mind.

Indeed, he got his degree here at Penn in, I believe, 1930, in engineering, but then over his career, went on to think about things related to the mind, and indeed wrote this book, *Consciousness, Matter, and Energy: The Emergence of Mind In Nature*, which treats these issues, and they're issues to which the scientific study of cognitive science speaks today.

And each year, we bring in a distinguished speaker, whose work relates to the mind and to consciousness. This year, we're fortunate to have with us, David Pesetsky from MIT.

And he'll be introduced by Julie Legate, who is an Associate Professor in the Department of Linguistics and Chair of our undergraduate major in Linguistics.

So before Julie comes up, I'd just like to present David, with this copy of Benjamin Pinkel's book. Here you are, and welcome to Penn. And so with that, I'll hand the microphone over to Julie.

MS. JULIE LEGATE: Thank you. It's my great pleasure today to introduce to you, David Pesetsky. David is one of the preeminent linguists of our time.

He's the Ferrari P. Ward Professor of Linguistics and Philosophy at MIT, a chair that was previously held by Nam Chomsky, Morris Halle, and Ken Hale.
He's a fellow at the American Association for the Advancement of Science, and a fellow at the Linguistic Society of America.

He also, at MIT, is the Cecil Green MacVicar Faculty Fellow in honor of his excellence in teaching and mentorship, something that I can personally attest to, as I had the good fortune to study under David at MIT. He was instrumental in forming my ideas on linguistics and syntactic theory.

Within linguistics, he's an extremely influential scholar. He's made important discoveries in syntax, in morphology, in language acquisition, across a broad range of topics, including questions, selectional properties of verbs, the structure of verb phrases, and argument structure, tense, and agreement, and case, and linearization, fundamental questions in linguistics.

His work focuses on discovering the laws of language that underlie the superficial diversity that we see among the world's languages. His interests and influence expand beyond linguistics, however. He played a fundamental role in initiating cross-disciplinary graduate training between linguistics, psychology, and cognitive science at MIT, starting in the early 1990s.

He's also studied how linguistic and psychological sciences can inform our understanding of how children learn to read, and inform educational policy thereof.

He's also an accomplished violinist. He is the principal second violin in the New Philharmonia Orchestra of Massachusetts.

And in accordance with his major theme of his work in discovering the fundamental relationships among seemingly disparate phenomena, today he will be discussing for us, the relationship between music and language. And so his title for us today is, “Language and Music: Same Structures, Different Building Blocks.”

Thank you.

[Applause]

MR. DAVID PESETSKY: Thank you very much. The mic is on? Thank you very much for a very generous introduction. Thank you very much as well, for this wonderful invitation to give this distinguished talk.

Thank you, Apple, for telling me that I have turned off my screensaver, but lying to me about it, evidently, as you just saw.
[Laughter]

MR. PESETSKY: I'm going to begin with a few remarks of humility and cautionary notes. I'm very, very honored to be here in a long line of extremely distinguished people, who have given these Pinkel lectures. It's an extraordinary honor to be asked to give it.

But my talk today is going to be quite different from the talks that you've had in previous years, at least from what I can judge from looking at some of the videos and some of the other material.

Previous Pinkel lecturers have been asked to come and talk about their day job. And if I were here doing that, I would be talking to you about the syntax of Russian numeral phrases, or hypothesis that laws that govern anaphora in language and laws that govern the assignment of case morphology to nouns might be deeply the same, or various other things that students of ours at MIT or I, myself, have done.

But I think because there's a particular body of work that I've begun over the last couple of years with a former graduate student of ours, Jonah Katz, that has some interdisciplinary interest above and beyond the other research that we do, that I was asked to come and talk about that today. But this is not my day job.

I am a linguist, and if I'm discussing syntax, I can be as wrong as anybody else. But I have some reasonable confidence that I at least know the likely areas where I'm going to be wrong, and I sort of feel on firm ground, when I talk about it.

I'm talking today about the relationship between language and music. And one side of the equation, I've done some thinking and I've done some work with my collaborator, but I'm definitely an amateur on the music side of the equation.

Penn, for example, has a very distinguished Music department and a distinguished history in music. I actually remember where I was in college when, in a class, they played me Rochberg's Third String Quartet. He was a professor here. It was a pivotal moment in my amateur musical life. But though I'm going to be talking about music theory, and I've talked about these topics to music audiences before, without being killed, I'm here alive, as you see me today, it will be at a lower level of competence than I can exhibit in linguistics. But I think that's sort of essential, if there's anything to the topic and to the ideas and to the questions that I want to discuss today.
At least at the beginning, this work is going to be done by people who are professionals on one side of the equation, but amateurs on the other side of the equation, because there isn't yet a field that will have professionals on both sides of the equation. And the hope of, more or less, speculative work of the sort that I'm going to talk about today is that we may look forward 20 years from now, to a time when there is such a field, and I won't be the logical person to invite to come and talk to you about language and music, their structures, and their building blocks.

This is joint work with a former graduate student of ours at MIT, who is now a post doc in the Phonetics Lab at Berkeley, Jonah Katz. Jonah was an undergraduate Music and Linguistics major at UMass Amherst, and joined our program with a continuing interest in some of the topics that we're going to discuss today.

And sometime during his second year in the program, I ran into him in the hall, and I asked if he was still interested in this because I had always been interested in these topics, and if he wanted to start a reading group with me. And so we met over a long period, and it was some of the most fulfilling discussions that I've had in my career. And one of the results is the body of ideas that I'm going to talk about today.

What is the topic of this talk? Well, there are two conflicting intuitions about language and music, that I think just about everybody who thinks about the topic has had.

One intuition that one has is that music and language are deeply different. And when you look carefully at what we know about the structures of both domains, the structure of music, and the structure of language, this notion is justified by a number of facts.

For example, I owe this observation to a talk that I heard by the phonetician Bob Ladd a few years ago; though all languages manipulate pitch in one way or another, no language makes use of an octave-based scale system, pitch collection of the form that all pitch-based musical systems appear to do.

Another maybe now slightly more controversial observation is that in many musical systems, at least a steady beat, an isochronic beat is a key component of musical rhythm, but is perhaps unknown. It's now apparently being questioned, but at least possibly, not present as a component of language.

More obviously though, music is felt often in some very abstract sense, to
have a meaning in some vague sense of the word. Music never has a meaning in the sense that semantics within linguistics discuss meaning. No musical phrase is true or false, for example, in a particular situation.

And every normally developing human can talk, but not every normally developing human can carry a tune, at least not with equal ability or play an instrument, and so on and so on. There are many differences.

Of course, there's an equally rich, more lofty, perhaps, tradition of viewing music and language as deeply similar. So first of all, language and music may be at least rhythmically organized in commensurable ways, so metered poetry or the very fact that music can have lyrics.

Language and music, in a sort of vague sense at least, and I'll argue in a less vague sense as well, are law-governed, complex systems that vary in some ways across cultures.

Both language and music have an emotive expressive function, something that I won't talk at all about today, but it's very real, of course. And every culture appears to have language and with some caveats that you don't always know where the music is in a culture, a point that Ian Cross makes, music also appears to be a property of every culture as far as we know.

And both abilities appear to be specific to humans, even though we know cockatoos can – if you go to YouTube, they can entrain, perhaps, to the rhythm of bad music. Both abilities develop without specific instruction, and they are at least supported by abilities detectable even in newborn infants, and so on here as well. So they're deeply similar and they're deeply different.

These conflicting intuitions correlate with two traditions in the discussion of language and music, so there are two distinct academic disciplines that study the formal structure of music and language. So on the music side, there's music theory, and on the language side, there's linguistic theory. And they differ in their history, they differ in their descriptive apparatus, they differ in their intellectual goals. So that's on the distinctive side.

On the similarity side, there's a vast speculative, informal literature, and sets of remarks by famous people, linking music to language. The most famously and notoriously, Leonard Bernstein's Harvard lectures from the 1970s, published as The Unanswered Question.

This has consequences for people who, in the cognitive sciences, think about these topics. So because the cognitive sciences have inherited the consensus view that the formal systems governing linguistic and musical structures are
distinct, the consensus view in the sense represented by the distinct academic traditions that study these two domains, most research out there that compares language and music concerns issues of use, rather than the formal abstract structure of the rule systems underlying music and language.

There's any number of articles, some by colleagues at MIT, on topics like, "are memory resources shared by the processing of music and language?" It turns out that the answer might be yes.

Is there a selective impairment after stroke, say, of musical and linguistic skills? The answer there turns out to be yes, as well.

And when it's discovered that music and language share some resource or property, it's taken as a contribution to the topic that's under the slogan of domain specificity. And so it's a possible indication that the resource or property that's been discovered is not specific to either language or music, but might be domain general, something that transcends both music and language.

And that's kind of the consensus view and kind of the consensus approach. What I'd like to do today is at least call into question that consensus.

What I'd like to argue is that when the form of the structures that organize language and music are compared in a particular way, the consensus view might be wrong. And I want to at least advance the strongest possible version of that conjecture. I want to suggest that the syntax in music and the syntax, the internal organization of language, are or might be formally identical.

I want to talk a little bit about some history of the discussion, and then I'm going to clarify what the claim is, that I'm making.

There's a little bit of history of this discussion that's relevant to what I'm going to talk about today. As I mentioned earlier, in 1973, Leonard Bernstein gave the annual Norton Lectures at Harvard, which were published a few years later as The Unanswered Question. They're on DVD. They're probably on YouTube. In which he suggested, after having read some books by Chomsky and gone and talked to people at MIT, that there might be deep similarities between what we had learned about language at that point and properties in music, including universality and internal organization.

Ray Jackendoff, a linguist who I think gave one of these Pinkel lectures a few years back, and a composer, a young composer at the time, Fred Lerdahl, attended these lectures. And the lectures, as they later wrote, suggested
that Bernstein was no scholar. They were informal. A lot of what he said, if you try to look at it carefully, doesn't quite add up. They were inspirational. They were insightful. They didn't really drive home their point with any consistency or accuracy.

Jackendoff wrote a review of Bernstein's book, in the journal *Language*, that sort of made that point. But it made another point, which is that it's an exciting idea that language and music might have something formally in common, and it's worth investigating in a systematic way, to see whether Bernstein's conjecture was correct.

Jackendoff and Lerdahl started working together, inspired by Bernstein's lectures, to try to produce a proposal that addressed Bernstein's questions and his ideas more systematically and seriously. And the result was a book published in 1983 called *The Generative Theory of Total Music*, “GTTM” to its friends. And this book is going to have a lot to do with what I discuss today.

A word or two about GTTM. As I said, there's a centuries-old tradition of investigation in musical structure. And Lerdahl and Jackendoff, much as Chomsky in linguistics, build on the centuries-old tradition of investigating linguistic structure. It didn't invent nouns and verbs. It built on what previous work on language had contributed.

GTTM assumed most of standard music theory. It assumed that lots of what had been developed over the years by smart people, thinking about how music was organized, was correct. They just wanted to add to it, re-characterize it, ask some new questions, and see if they could find some new answers about it.

In particular, they identified one central figure in this tradition, the Austrian theorist Heinrich Schenker, who they saw as prefiguring a kind of Chomsky in their generative investigation of music, in ways that were very important to them, that we'll be talking about. So that is a little bit of background to GTTM, which is, in turn, background to this talk.

The starting point of Jonah and my work is actually, GTTM, on the grounds that, as far as we know, it's still the most complete and insightful formal account of the structure of Western tonal musical structure.

What we'd like to argue is that there is something deeply the same about music and language, which comes out of the GTTM discussion. What we'd like to suggest is that when the Lerdahl and Jackendoff GTTM model is reworked in a form that's commensurate with modern linguistic theory and compatible with its goals, the syntactic structures, the formal structures of both music
and language are at least in part built by exactly the same rule, the rule that linguists have recently taken to call, “Merge.” I'll tell you what Merge is shortly.

Another point that Jonah and I argue for, that I will be only alluding to today, and not dealing with seriously, is that the structures produced by this rule, Merge, and I'll teach you what it is, interact with computations of other components of the mind, in a similar fashion in both music and language.

The information flow and overall architecture of musical and linguistic systems may have something deeply in common as well, and that's a point that we developed further in the paper that you can read on-line, that I won't be discussing at any length today.

I am going to argue that the core rule that puts musical structure and linguistic structure together is identical in the two domains.

At the same time, music is not language and language is not music, so something is not the same. What's not the same?

What I'd like to suggest is that what's not the same are the building blocks that go into this rule, Merge. So another version of this talk, we sometimes call it, “Music and Language: Different Ingredients, Same Recipe.”

So the building blocks of linguistic syntaxes, we'll see, are words or word parts, morphemes, stems, listed in a mental database called the lexicon. The members of this database are called, lexical items. They are arbitrary pairings of sound and meaning, in the case of sign systems, arbitrary pairings of gesture and meaning.

And lexical items and the lexicon have no obvious counterpart in music. Music doesn't use a lexicon in this sense, there are no pairings of sound and music that go into musical structure.

The building blocks of total musical structure concern pitch class, what degree of a pitch collection, a scale, you are, and chord quality. And these, likewise, have no linguistic analogues either.

I already mentioned that as far as we know, no human language divides pitch space the way music does. There's no notion of optimal prevalence. There's no notion of scale system. Linguistic total systems are about high and low, and rising and falling, and so forth. But they're not about pitch class, and they're certainly not about chord quality.
The basic building blocks of music and language are clearly distinct. In a nutshell, what I'd like to suggest is that what language and music have in common is clearly not their building blocks, but what they do with them.

Jonah and I state this as what we call, “the identity thesis for language and music.” All formal differences between language and music are a consequence of differences in their fundamental building blocks, arbitrary pairings of sound and meaning in the case of spoken language, pitch class, and pitch class combinations in the case of music.

We'd like to conjecture that in all other respects, language and music are identical. Our conclusion, then, if you care about the slogan of domain specificity, is that, in many respects at least, music and language count as the same cognitive domain. In some respects, they're different.

We're going to take as our starting point, Lerdahl and Jackendoff's ideas in that book that I showed you, called GTTM. But actually, Lerdahl and Jackendoff discuss the family of ideas that would include the “identity thesis” and reject it. They state, in several places in their book, the conclusion that the generative music theory that they have developed does not look like generative linguistics. It does not look like the linguistics that, say, Jackendoff, himself, has participated in as a researcher in language.

We'd like to respond to this in several different ways. First of all, it turns out that in a number of technical respects, GTTM was a decade ahead of developments in linguistics. That is, several aspects of musical structure, as they discussed it in their book, resembled current ideas in linguistics much more closely than they resemble ideas that were current in the late 1970s, when they were writing their book.

We'd also like to note that despite the fact that the G in GTTM stands for “generative,” which is a term from linguistics meaning the precise development of formal models, the actual goals of GTTM, if you read what they're doing and you pay attention to how they do it, differ from the goals of linguistics as, say, Jackendoff himself practices it.

This difference, in turn, affected how the model was presented and which aspects they chose to emphasize. So GTTM was explicitly concerned with the following very important question: given a piece of music, and a skilled listener who knows the idiom in which that music was composed, what analysis does that skilled listener assign to it?

It was far less concerned with the more abstract question: what's the logical structure of the rule system that licenses the analyses provided by listeners
for the pieces that they hear?

And it's that other question that would be more analogous to those that are at the heart of most work in generative linguistics, which is the study of rule systems for language. But that's not a value judgment.

Given a rule system for English, how does a speaker hearing in English text, for example, parse it and assign meaning to it, in real time? That's an important question that many people work on. But it's not the question that's central to the enterprise that's normally called generative linguistics, while the corresponding question in music is what GTTM took up.

So that actually, inevitably, colors the look of the proposal that GTTM made for music, and helps make it look more different from work in generative linguistics than maybe it needed to be.

We're building very much on GTTM, so you might ask, am I just giving you a book report on GTTM? Is that worth a Pinkel lecture?

[Laughter]

MR. PESETSKY: Well, maybe we are, but I think it's not just a report on GTTM, that we're going to get.

What we'd like to propose is a realignment of, essentially, the GTTM model, and I'll show you what you need to know about it shortly, with the kind of linguistics that I do in my day job.

When we do that, some similarities, and the differences as well, between language and music, if we're right, can be examined, perhaps, more meaningfully and more productively, than GTTM's presentation allows. And if we're correct, this realignment supports, or at least supports the interest of, the identity thesis.

Methodologically, note that we should be very happy to start with the identity thesis, and then learn ways in which it's wrong. But it's going to be most instructive to start, as is always the case, with the notion that two dissimilar things really are identical, and then learn where the differences lie, than to start with the presupposition that they're completely different, especially when there's some reasons to have a hunch that the similarities are real. So that's what we're trying to do.

Let's talk a little bit about language. My understanding is that I'm talking to an interdisciplinary audience. Not all of you are linguists, so I'm going to give
you a little bit of elementary linguistics because it's important to the music that we're going to discuss.

If you have a toy sentence in English like, “The girl will read the book,” there's good reason, you can read hundreds of pages of arguments for this and arguments about details in the linguistic literature, to think that the structure for a sentence of this sort can be represented as a hierarchical tree of the sort that you see on this slide, where the nodes of the tree indicate for us, the constituents or the sub-groupings of the sentence that are relevant, both to other aspects of the syntax, and also, to the semantic interpretation of the sentence.

“D” here stands for determiner. That's the part of speech in fancy linguistics that “the” belongs to. “T” stands for tense. “N” self-evidently stands for noun. “V” stands for verb. And the “P,” as in “TP” and “VP,” for phrase, verb phrase, tense phrase, and so on.

These structures can be characterized as the product of a rule that takes, as its initial input, lexical items. So we have a lexicon in our heads, lots of theories, lots of debate about that. Penn is a hotbed of work in one such theory.

I am not going to give you anything sophisticated here, so let's blur the distinction between morphemes and words, just for the sake of today's discussion. And imagine that the lexicon consists of words like, “the,” and “girl,” and “will,” and “read a book,” and consists of a body of information about each of these items, for example, the part of speech that each item belongs to. So, “the” is a determiner, “girl” is a noun, “will” is a tense, and so on and so on.

Some information about the meaning. “The” means something like, it's old information, we've been discussing the girl before. The hearer can reasonably know which girl we have in mind. “Girl” is something like, a young, Cuban female. “Book” is something like, data downloaded to a Kindle. Nothing but up-to-date here. And so on, as you see up there.

We have a large lexicon, probably more interestingly structured than you're seeing on this slide, in our heads. That's the basic building blocks that go into the structure of a sentence of human language.

And the rule that puts these building blocks together has been named by Chomsky, “Merge,” and it's a convenient name, so it's what I'm going to use here.
So it works something like this. You construct a phrase. What you do is, you form a set from two elements, we'll call them “X” and “Y.” And the key thing is that X and Y can be either lexical items, or it can be a set that we previously formed by applying that rule. That's the recursive step.

Once we've formed a set from two elements, X and Y, where either or both of them are lexical items, or a set that this rule previously formed, we choose either X or Y as a special member of the set we just formed, as a special member of that phrase called “the head.” The head gives the phrase, its name and helps determine the properties of a phrase. The properties of a phrase are determined by the sub-constituent that is its head.

For example, we could take from the lexicon the determiner “the,” and the noun “book,” and we can put them together and form the set, “the book.” And we can decide, and this is perhaps slightly controversial, there's a debate as to whether “the” or “book” is the head of the phrase. I'll follow most linguists today, in assuming “the” is.

And so we form the set, “the book,” of which the determiner “the” is the head. The result of the nominal phrase is a DP, a determiner phrase, which we can represent in a tree like that.

Then we can go to the lexicon again and find the verb “read,” and merge “read” with the set, with the phrase we just created, forming the larger set, whose first member is the verb “read,” and whose second member is the set “the book,” that we just formed by Merge, which we can represent in a tree structure looking like that.

And there's evidence from linguistics that the head of that constituent, that phrase that we just formed, is the verb. It's the properties of the verb that give the phrase consisting of the verb and its direct object its properties in turn, which is why we call it a verb phrase.

The rule is recursive because one of the elements that entered into the set is of the set, the Merge, that was formed by a previous application of the same rule.

And we can do that, once again, with the tense element “will,” giving us “will read the book,” as you see in this slide. And then meanwhile, in a galaxy far, far away, we could have started the process separately with two other lexical items, for example, “the” and “girl.” We can build a DP of the form “the girl,” and then we can go and take “the girl” and merge it with the phrase “will read the book,” giving us the tree that I showed you a few slides ago.
That's how recursive Merge works. “The” merged with “book.” “Read” merged with the set “the book.” “Will” merged with the set “read the book.” “The” merged with “girl.” And the last step in forming the structure is to take two complex objects, the set “the girl” and the set “will read the book,” and merge them together.

And again, there are reasons in the linguistic literature for assuming that the head of the sentence that we just formed is, in fact, the tense of the sentence. That's not important to us here, but it is a fact.

And, of course, we can keep going. We could add what non-linguists might call a subordinated conjunction, what linguists would call a complementizer, hence, the letter “C.” The word “that” converged with “the girl will read the book.” “that the girl will read the book.” And then that, in turn, can merge with a higher verb, and we're off and running for a sentence like, “Mary believes that the girl will read the book.”

This is building a complex, potentially infinitely, complex structure by recursive application of the rule Merge. And at each stage, we choose a head. The head of “read the book” is the verb “read,” as we saw. The head of “will read the book” is the tense element “will,” and so on and so on.

Heads in these phrases matter deeply to linguistics. They are kind of what syntax is all about. And semantics cares about it as well.

For example, heads care about the heads of things they merge with. The English word “that,” the subordinating conjunction, the complementizer of “that,” merges only with a phrase that's headed by finite “T:” “that Mary will read the book.”

There's another complementizer, “for,” as in, “I would prefer for Mary to read the book,” that merges only with the phrase headed by “to,” by the infinitival T. And the lexical item chosen for T, if it's “will” versus “has” versus “be,” “be reading the book,” determines the suffixes, the morphology on the verb. Is it a part as simple as an infinitive?

And verbs care about what they merge with as well. That's what meant by being a transitive or an intransitive verb.

Heads also matter across linguistically, in determining word order. So the relative order of heads and the elements that they merged with vary systematically across languages.

English is a head-initial language. The head of a phrase systematically
precedes its sister, the element it merged with. So “that Mary will read books,” “read books,” the head of the verb phrase, precedes the non-head.

But there are head-final languages like Japanese, in which the head of a phrase systematically follows its sister. “Mary books read will that believe.” That's Japanese. “Sue Mary books read will that believes.” The verb follows its object. The tense follows the verb phrase. The complementizer, the subordinating conjunction, follows the sentence that it attaches to.

There are mixed languages as well. For example, if you look at subordinate clauses in German, which are where you can really learn how German word order works, you'll see that it's head-final off the verb phrase. “John said that Mary books read will.” And that the tense element comes last in the sentence, so it's head-final, this is controversial, arguably, head-final at the tense level as well. But complementizers, subordinating conjunctions, precede their complements, precede the things they merge with. So it's head-first at the complementizer level.

In at least one dialect of Afrikaans, the dramatic language of South Africa, it's like German in that the verb follows its object, but like English in that tense precedes the verb phrase. So it's head-last of the verb phrase, head-first at the sentence level, and it's also like German, head-first at the complementizer level, at the subordinating conjunction level.

You have languages that are mixed in being head-last of the verb phrase and head-first higher up, just like you have languages that are head-first everywhere or head-last everywhere.

You could easily imagine other sorts of mixed languages as well. For example, in literary Martian, the verb precedes its object. So you “read books,” you don't “books read.” And tense precedes the verb, so you “will read books,” you don't “read books will.”

But the subordinating conjunction follows the sentence. So in literary Martian, “John believes Mary will read books that.” Or Old High Venusian has a head-initial verb. So you “read books,” but the tense comes at the end. “Read books will.” And the subordinating conjunction comes at the end.

So as you may guess, I hope you've guessed, literary Martian at least does not exist on this planet and this species, as far as we know, nor does Old High Venusian exist on this planet and this species. In fact, there's a universal, apparently, a universal property of languages.

Some issues arise with things like, Mandarin Chinese question particles. But
apparently, universal property of languages, that within a single clause, if a phrase is head-initial, like the verb, for example, precedes the object, the phrase that immediately contains it must also be head-initial. So going from the bottom of the clause to the top, if you're head-initial anywhere downstairs, you have to be head-initial upstairs.

On the other hand, if you're head-final downstairs, as in, say, German, it's perfectly okay to be head-final downstairs, head-final downstairs, and then head-initial upstairs. What you can't be is head-initial downstairs and head-initial upstairs.

These kinds of languages seem not to exist. At a body of work spawned by a paper by Andrus Holdberg and developed in collaboration with other colleagues in English, it has suggested, in surveying a large number of languages, that this is a general property of human languages. They call it the final-over-final constraint.

Just to keep your disappointment level low, I'm not going to demonstrate the final-over-final constraint in music. That'll be an interesting question. But I'm introducing it here, as a way of showing you, the importance of the notion “head of a phrase,” in a hierarchical structure in language.

Let's talk about music a bit. GTTM presents a sort of architecture for their formal model for music that sort of looks like this, kind of like the Philadelphia subway map or something. And crucially, it doesn't look at all like comparable boxologies, comparable architectures, that are typically proposed for language.

Nonetheless, what Jonah and I have argued is that deeper similarities emerge if we actually focus on the function and interaction of specific components.

Let's begin with a component of this model that Lerdahl and Jackendoff in GTTM called, “prolongational reduction.” Lerdahl and Jackendoff, developing Schenker's more informal ideas from earlier in the 20th Century, argued that underlying total musical structure, the harmony of a piece of music is organized much like the linguistic trees that we've just seen, into a binary branching, headed tree which they call prolongational reduction or PR.

And the heads of this tree are chosen so as to maximize distance between sister nodes in the musical surface. So the distance between two chords is measured in the pitch space that takes account of the notion key, and the relative stability of connections is determined in ways they talk about by the distance from each other.
Consider, for example, a quite obnoxious toy melody that I composed, not because it was obnoxious, but because it was short and simple and stupid. That sounds like this, if the audio works.

[Background Music]

MR. PESETSKY: Not much to it. It's not a great composition, with music theory. The people in the audience, don't check my voice leading. I'm just being stupid here.

And the numbers underneath it are just numbering the chords from left to right. They have no significance. But in the GTTM model, the prolongational reduction structure for this melody might look something like the tree that I've drawn over this structure.

This tree has visual properties that are unfamiliar to linguists. So familiar to linguists are the fact that the tree has nodes, that it's an acyclic, directed graph. Each node has only one mother.

But you'll notice that the lines connecting various nodes to other nodes either involve line segments, or are they straight, or they involve an angle somewhere. That's just GTTM's way of notating head. This little bit over here, five in GTTM's notation, is the head of the constituent formed by putting it together with four, and four is the non-head.

We could rewrite this tree in the kind of notation that linguists would be familiar with. Again, the numbers simply refer to the sequence of chords from left to right. And so for a linguist, we would simply say: this is sort of the phrase, whose daughter nodes are a five and a four, and the five is chosen as the head.

The idea would be that this is the stronger, more stable note of this pair, but that these two chords belong together structurally. And, of course, this kind of structure then, in its abstract properties, is reminiscent of, in fact, formally identical to, what linguists propose for syntactic structure.

The idea from GTTM is that in a structure of the sort that you see in the slide, the head of XY, the less tense member of the pair is the chord that's harmonically closer to Z. That's how they work it all out.

And their idea is that the tension associated with an event depends in part upon its structural context in the tree, and that these patterns – this is the key thing – are not explainable purely in terms of linear adjacency, purely in terms of left and right. And there are experimental results that support the
structures that they're assigned.

So consider the obnoxious little tune that I just played for you and the proposal as to what the structure is. I think you can easily see that these structures can be viewed as arising from recursive Merge. It's a Merge because the structure is hierarchical. It's recursive because the rules governing what merges with what, and the choice of head, appear to be the same throughout the structure. There's the same sort of principle underlying all of it.

You can imagine merging chord one with chord two, forming the set one-two, of which the head is one, going on to chord three, and so on, as you see in this slide. Ultimately, resulting in the tree that one would assign to this in the GTTM model.

Now, this is sort of easy-to-come-by evidence for this kind of organization. Because of principles that tell you what goes with what, and what's the head, are identical throughout the structure, the higher levels and at the lower levels, if you were to perform the structure in a way that substitutes the head of any constituent for the full set of terminals of that constituent, the full set of members of that constituent, let's call that a reduction, the result will not only continue to sound like a piece of music, but will actually sound like a simplified version of the same music. It's sort of a principle behind the simplest case of theme and variations.

If you have the obnoxious little tune – I'm sorry. That wasn't my intention. My intention was to play the – let me go back or let me not go back. Whoa. What's happening? The goal here is...

[Background Music]

MR. PESETSKY: If you have a little tune here, and let me see if I can restore what we're doing here, and you delete the non-heads of the lower, of this phrase and of that phrase, the result will, if I can manage to click it, sound like this. I don't know why it's not doing its job. It should sound like this.

[Background Music]

MR. PESETSKY: It's just a simpler version of the same melody. And then if you just delete the non-heads of the higher constituents, the result, if I can make the machine cooperate again, should sound like this.

[Background Music]
MR. PESETSKY: So what we're doing is, we're just getting – at each successive level, going up the tree, we're just getting rid of the non-heads. Ultimately, if we continue the process and if my machine cooperates, we have this.

[Background Music]

MR. PESETSKY: And it's that intuition, simply, that in the sequence of musical events that made up my obnoxious little melody, some events were more important than others, that they are actually arrayed in a hierarchy of importance, where the important member of each pair is what we're calling the head.

That's sort of validated by the very simple musical intuition that there are notes you can remove from the structure, that in some obvious sense, preserve not only the musicality of the phrase, such as it is, but our sense of the actual identity of the melody.

And random deletion won't do the trick. If I take out the heads of constituents, for example, while maintaining their non-heads, you might end up with the following sublime piece of music.

[Background Music]

MR. PESETSKY: I mean, you can decide for yourself, whether that's something you want on your iPod, but be that as it may, it's in no way the same melody anymore. It's not recognizable as the same melody, the way the other variants were.

What we're looking at crucially here, and that's of interest to a linguist, is nonlocal interactions among nonadjacent elements. The fact that reductions, where we're taking out some middle notes, continue to sound like a version of the same piece illustrates the possibility of nonlocal interactions that follow from a hierarchical structure that's also headed.

We could call this “Schenker's Generalization,” filtered through GTTM, that two chords interact harmonically, so long as they are the heads of phrases that merged with each other. And that what's crucial and crucially interesting is that the heads themselves need not be adjacent to each other. They can actually be separated by all sorts of florid stuff, if you want.

We have a postdoc now at MIT under the Intelligence Initiative that last year's Pinkel lecturer, Josh Tennenhouse and Tommy Pojow have organized. And Martin Rohrmeier, who's doing work in a rather similar vein to what we're talking about, he gives the nice, actually, simpler example of a
standard conventional passage of this sort.

[Background Music]

MR. PESETSKY: I apologize for the bad audio, but just another example of the same sort of thing.

Take out the second of these chords, and you still sort of have music, and it still sort of sounds right.

[Background Music]

MR. PESETSKY: Take out the second and third, and it's kind of short, but still...

[Background Music]

MR. PESETSKY: But if you take out, say, the third, without taking out the second.

[Background Music]

MR. PESETSKY: Again, it doesn't kill you, but it's not satisfying in the same way. Something has gone wrong with that example, that was not going wrong with the previous examples. Likewise, if you take out the...

[Background Music]

MR. PESETSKY: Again, it doesn't kill you, but there's something off there. And what Rohrmeier argues is that what's going on is exactly the same principle. Syntacticians who remember transformational grammar will know that what we're talking about is factorization.

You can play the whole sequence, or you can factor it as this chord and the head of that one, and then that one, and that will work. You can play the C, the G, and the C. I suppose you could also play just the C. And that's all fine. Random deletions don't work. There is a structure underlying musical sequences of this sort. Nonlocal interactions.

Remember, nonlocal interactions is actually the key to a hierarchical head in musical structure. For example, in English, a tense must merge with a verb phrase. “Mary will be happy,” or something. And it can't merge with something that doesn't have a verb in it. “Mary will almost certainly happy.” That's not English.

But the tense and the verb of the verb phrase don't have to be next to each
other. “Mary will almost certainly read the book.” The verb that the tense needs is several words away. It's okay because it's heading the phrase that's actually the sister of tense.

Recall our proposal, the identity thesis. Formal differences between language and music are a consequence of the pairing of differences in their fundamental building blocks. In all other respects, language and music are identical.

So language – what are the differences in the building blocks? Well, language makes use of a lexicon, whose items pair sound and meaning. Music does not. Music makes use of positions in a scale system, organized throughout an octave. Language, it does not.

But what we're suggesting is that the combinatorics are identical. The distinct building blocks are combined by the same rule, Merge, forming identical binary branching, hierarchical structures, including the notion of head.

Then you might go on to ask, does the result generalize? Linguists who work on Merge, who work on syntactic structure, note more varieties of Merge than I've been talking about so far.

So are all types of Merge, all types of structure building that are found in language, are they also found in music? If the identity thesis is correct, all things being equal, the answer should be “yes.”

What about a species of Merge that linguists these days sometimes call “Internal Merge,” otherwise known as syntactic movement? Is that found in music?

First, for the non-linguists, let me give a brief explanation of what that is. Given straightforward expectations about syntactic structure, like a verb should merge with its arguments, and given straightforward ideas about how linear order is established, like when two elements come together, the right edge of one should align with the left edge of the other.

We don't expect, and it's interesting to discover, that in many cases, the otherwise predictable relation between hierarchical syntactic structure and the phonological surface breaks down.

For example, we've been talking about tense and the verb sort of being in a relation with one another, and the verb and its complement being in a relation with one another, that leads us to expect, all things being equal,
that the verb and the complement should be next to each other.

But in French, as is well-known, in a finite clause, the verb is actually where the tense should be. The verb is not necessarily next to its direct object. Treating “pas” as the real negation in French, “The girl will not buy the book” comes out as “The girl will buy not the book.” And why is that? That’s a ubiquitous phenomenon in the languages of the world, or the common phenomenon.

The hallmarks of the phenomenon that we're seeing in French, is that the finite verb is pronounced adjacent to tense. That is, the tense here is that – probably, the dash should be in a different place – but the tense here is the suffix on “buy” between, you know, there should be a dash between the T and the E, that has to be adjacent to the verb “buy.” They form a single word. Finite verb is pronounced adjacent to tense, and is tightly coupled with it. The two of them form a word.

Yet, at the same time, the verb has its expected syntactic dependence. The direct object is there. It's just, the verb isn't sitting next to it, as we expect. The verb is sort of seemingly dislocated to T, somehow, but it's still heading its own verb phrase. There's still a verb phrase there.

One way in which linguists analyze this type of phenomenon, and there are instances of it in English as well, is that in a French sentence like, “the girl will buy not the book,” the verb, “acheter,” “buy,” did actually merge with its direct object, as expected by the normal laws that associate meaning with structure. It's just that they merged a second time with T, as illustrated in this structure.

Here's the verb, and this should be D-P, taking “the book as its direct object, and the verb merged with this D-P, “the book,” and then merged a second time with the tense of the sentence. It's literally, in this structure, occupying two distinct positions, and it's getting pronounced in the higher of those two positions in the structure.

This is a phenomenon called “head movement,” here of the verb to tense, and it's very common in languages in the world, and English used to do it as well. It no longer in general does.

The phenomenon has as its hallmarks, adjacency, the verb and tense are next to each other, but there's still a verb phrase.

Another thing that we observe in French is that the movement is obligatory; that is, the movement of the final verb to tense satisfies some need of an
element in this structure, probably, tense, since non-finite verbs, infinitives in French, don't raise to T. So “ne pas acheter le livre,” the finitive verb follows the “pas;” it doesn't precede it.

The tense-verb relationship that ends up with the verb moving, merging a second time to tense, in some way alters the properties of tense. That is, tense had a need, and the need gets satisfied by the verb moving to it. And so it satisfies the needs in some way.

And we've already observed that another hallmark of the process is that, once the verb is in T, the verb and the tense are tightly coupled. They function as a single word, in a general characteristic of head movement.

You can ask in music, are there also cases in which an otherwise expected relation between structure and linear order might break down? That is, what should we look for, if music, for example, shows head movement, like the French example that we saw?

What we would expect would be, for example, in a total passage, that some chord X has to be performed adjacent to some other chord Y. While at the same time, X has a normal set of syntactic dependence of its own, linearized normally, and thus, apparently, also heads its own phrase.

We would also look to obligatoriness. The X that moved to Y might be obligatory, arguably satisfying some need of Y, required in order for the, as linguists say, derivation to succeed in order for the structure formed by Merge, to satisfy the requirements of the system.

And we might also look for X and Y to be tightly coupled together, so that for other purposes, they function as an indivisible unit.

What Jonah and I have argued, have suggested is that in total music there is a place where we find these properties. It's what’s called a cadence, in particular, a perfect cadence, a perfect authentic cadence. It's a sequence of two chords, a dominant and a tonic, represented here by delta and tau. But a tonic is a chord built on the first note of the scale, called the tonic in music theory. And the dominant delta is a major triad, most often built on the pitch five scale steps higher than the tonic that's called the dominant in music.

And the cadence is a big deal. It's a major structural signpost that establishes the key of a passage and marks the ends of sections of musical pieces in standard total music.

What we find interesting is that a cadence has the following additional
properties that we think are remarkable: the first element of the cadence, the dominant, the delta, must be adjacent to the second element, the tonic. While at the same time, the dominant has a normal set of syntactic dependents, linearized normally, and thus, apparently, also heads its own harmonic phrase, called a delta P.

In particular, the presence of at least one dependent of the dominant called the subdominant, represented by upsilon for the U, and subdominant, most often built on scale degree four or two, is exceedingly common and maybe even obligatory in certain styles.

Consider, for example, the obnoxious melody, and listen carefully to the last two notes, which, if I can get it to play, have been made louder here.

[Background Music]

MR. PESETSKY: It's those last two pitches, the last two chords that are the cadence. That's circled here.

You can imagine a formal proposal that looks like the following. That the obligatoriness or the near obligatoriness in standard pieces of the perfect cadence, we know that it's kind of preconditioned for satisfactorily having established the key as the key of the relevant tonic.

You can imagine that what's going on here is that the tonic is actually marked as a tonic, licensed as a tonic and the key by the chord that precedes it, by the dominant, and that this feature plays a role in the structural condition on key establishment.

You can say, if a syntactic phrase is in the key of tau that's because it's contained inside a larger phrase, whose head has been marked as a tonic instance of tau.

And if total pieces in the relevant idiom must be in a particular key, it follows then that a piece must be headed by a tonic marked chord. It's a fact that standard, sort of, boring pieces, and in turn, some interesting ones too, are right-headed, so the cadence ends up being at the end of the relevant passage.

And in GTTM, in the technicalities of GTTM, in various ways that I won't go into here, the dominant and the tonic, when they're part of a cadence, actually behave as a unit. They have another system that I'll allude to briefly, where rhythm plays a role.
And even in passages where the dominant and the cadence, due to rhythmic subordination, should not be attached very high in some structures that they talk about, in a cadence, it has to be. They call this property, conditional retention, and I won't go into it. But at least in their system, there's evidence for what I've called tight coupling, as well.

Here's the question that I think should be asked and that, perhaps, has an interesting answer here. Why is adjacency, string adjacency, linear adjacency between the dominant and the tonic, a condition for the dominant to mark the tonic as a tonic, for delta to mark tau as a tonic? Why do they have to be right next to each other?

Why is it not sufficient simply for a phrase headed by the dominant, to merge with another phrase headed by the tonic, even if the actual heads, the dominant and the tonic chords, might not be next to each other? Why do they actually have to be adjacent?

This is actually very surprising, given the Schenkerian foundations of the rest of the musical syntax, because the whole point of which is that harmonic relations may be long distance. That's the essence of all that demonstration of, “take away some notes, it sounds like the same piece, take away the wrong notes, it doesn't anymore.”

All other musical relations between chords appear to require adjacency between the larger units that they head, non-adjacency between the heads themselves.

For example, you would take the C major Prelude for “The Well-Tempered Clavier,” I mean, you all know it, but I forget whether I have here, just the chords, or a Sibelius music program playing.

[Background Music]

MR. PESETSKY: That's not a particularly inspired performance that the Sibelius Music Program would play to - - - I make no claims for the performance that you just heard. I wanted you to hear it, in part, to clear your head of my obnoxious melody.

This is Lerdahl and Jackendoff's analysis of it, and there are some criticisms of it out there. But the thing I wanted to point out is that, here's a beautiful and extremely prominent, dominant chord, which is a sister of the final tonic.

Why isn't that enough to make the cadence happen? Why does the second to the last chord actually have to be another dominant chord? That's not how
the rest of the system works.

If total harmony is all about head-to-head relations that can be long distance, why isn't that enough to create the sensation of cadence? Why is this a general property of pieces in this idiom?

Jean-Philippe Rameau, who was a great theorist, as well as everything else, sort of notes that a perfect cadence has a certain way of ending a strain which is so satisfying that we desire nothing further after it. But why does it have to be at the end? Why does the dominant have to be the second to last chord?

If music is all about heads, why isn't Chord 25 sufficient to produce a sensation of cadence, but the final Chord 36, why must delta be adjacent to tau?

You could ask the same question in French. You could say, if language is all about heads, why isn't the fact that tense merges with a verb phrase enough to produce a sensation of grammaticality in a French finite clause? Why must the verb be adjacent to T?

So one possibility is that in this place, where we observe a convergence of a need of one chord being satisfied by another chord and an adjacency requirement, what we might be looking at in musical structure is a musical counterpart to the movement of one head to another that we find in language. That is, the idea would be that, let's say, delta, the dominant chord, merges with the subdominant chord that precedes it and any other dependents, and then merges a second time with the head, T.

In my obnoxious melody, the second to last chord, I'll just move it this way, the second to last chord, which could, in theory, have been anywhere, ends up being right next to the last chord because it underwent head movement to it.

Likewise, at the very end of the Bach C Major Prelude, the dominant chord that, in theory, could have initially merged anywhere, ends up having to be adjacent and tightly coupled to the final chord, because it underwent internal Merge that has head movement to it.

That's actually the main, novel point of interest that I want to present. My understanding is that in music theory, in the Schenkerian tradition, oddly enough, this question is not asked. That is, in the essence of Schenker, that relations can be long distance, but why some of them might not have to be, is not an asked question, so we want to ask it.
At the same time, you might find the argument weak, and you're not necessarily evil for doing so. The musical argument for head movement is different and weaker in some ways from the linguistic argument.

First of all, I think there's some reason for this. Because there's no lexicon in music and no piece-wide setting of head first or head last, unfortunately, in music, unlike in language, it's not possible, or at least, we haven't been able, to pinpoint the position where we might have expected to hear delta, if it had not undergone conventional head movement.

This contrasts with languages like French, where we can use notions like, direct object of, and the semantics that goes with it, to localize the position of the verb in examples where the verb has undergone movement.

For this reason, the evidence for head movement in music, if we're ever going to have it, must be more indirect than corresponding evidence in language, and I think there's nothing that we can do about that. We cannot have an argument of the form; we know the dominant must have started here, and look, we're actually hearing it 17 chords to the right.

Instead, what we can say is, look, we have reason to think that the dominant should be able to be any number of chords to the left of the tonic. If all we have is Schenkerian-headed, binary branch instructions, why is it that the only place we can hear it is right next to the final chord? I don't think we can logically do better than that.

To summarize the evidence, the presence of an absolute string adjacency condition on a pair of heads, linked to the satisfaction of a need on the higher head, in a domain of musical grammar where all other interactions require only sisterhood between phrases, that is, larger units headed by, say, two elements, X and Y, not linear adjacency between X and Y, itself, may argue for the existence of head movement-type processes in music as well.

Just a few more remarks and then I'll end. So harmonic structure is not sufficient to account for all intuitions about production. Symmetrical strengths, spacing among events, parallelism, all play some role in explaining theme and variations judgments. For this reason, if you know GTTM, they proposed a second kind of headed structure that coexists with harmonic structure that they call “time span reduction.”

In the paper that Jonah and I wrote, we observe, actually following a hunch of Lerdahl and Jackendoff themselves, that this second kind of structure is actually formally identical in some ways, to popular representations of
prosody, the phrasing, intonational phrasing in language, for example, Lisa Salkirk's proposals, and represents similar notions of prominence of phrasing.

We argue in the paper that the relation between the kinds of trees that I've been talking about here and this other structure, which is what Lerdahl and his subsequent work is focused on, is perhaps the musical counterpart of the relation between linguistics, syntax, and prosody, so it's musical prosody, much as Lerdahl and Jackendoff themselves suggest.

I'm running short of time, so I'll just flash this slide in front of you.

We also have a treatment of something called modulation, the circumstance in which recursively, a passage in an alien key can be imbedded inside a passage that's in some other key.

And we present a formal argument that at least the logical structure of the mechanisms to handle phenomenon of applied dominance and modulation resembles, in its interaction with the syntactic component that we've been discussing here today, the relationship between interpretive components in language, such as semantics. I don't mean lexical semantics, but just interpretive components and the syntactic component. But I won't discuss that in any detail right now.

What we think might be true conjecturally is that when we view GTTM's discoveries and proposals and the evidence underlying it, through the prism of modern linguistic theory, some formal similarities emerge between music and language.

Contrary to GTTM's own view of their own discoveries, their own discoveries perhaps do indeed look much like generative linguistics, contrary to their own conclusions about their own work.

If this is right, it at least suggests that there might be some support for the identity thesis, that all formal differences between language and music are a consequence of differences in the fundamental building blocks. The differences are real. But that in all other respects, formal respects, internal architecture of the system, language and music are identical.

Highly unlikely that this is going to turn out to be true, but who knows, right? But at the very least, what I would like to convince you of is that the identity thesis is at least fruitful as a strategic starting point for research on musical linguistic grammars.
And at least the sort of linguist's eye on musical structure, the program that was initiated by Lerdahl and Jackendoff, might also yield interesting results in the understanding of music, itself.

There's a million questions that you could ask me. And I had a wonderful lunch with some of the undergraduate majors in cognitive science and linguistics and music a while ago, and most of the interesting questions that you could ask me are going to get the answer, “I don't know.” So we can cut the question period short right here.

For example, in language, we have support for the views that we have from relatively detailed and increasingly detailed and deep analyses of a variety of systems.

What about in music? Cultures whose musical systems differ from that in western total music. Do they differ from one another in ways reminiscent of language variation? What about headiness and so forth?

I think we don't know the answers to these questions yet, but I think if our identity thesis is fruitful, it might inspire people to find the answers to this.

Why are languages head-initial or head-final? But the trees that I've been showing you are, to some extent, headed every which way. This is actually an old slide, and the answer is, “good question.”

Martin Rohrmeier, who I mentioned earlier, at MIT, actually has some ideas on this point, that conceivably, the musical structures in question are not headed every which way, but are actually uniformly right-headed, with two particular kinds of exceptions. So this might be a question that will have an answer.

Lots of other questions that you will probably ask me. I think I'll just flash this slide to you.

I'll just sort of conclude by saying, again, that what I've suggested here is that the identity thesis for language and music, if far from proven by this talk, at least my hope is that it's inspired some interest, and that it inspired somebody out there to go and answer some of the unanswered questions, show that some of the proposals that I've made are wrong or right or need to be modified, the kind of work that I know and love in my day job in theoretical linguistics.

And I want to thank you again for coming to my talk today. Thank you.
[Applause]

*Note: the video portion of the 2013 Pinkel Endowed Lecture ends here.*

MR. PESETSKY: Questions? Yeah.

FEMALE VOICE: Going back to the very beginning of the talk...

[Background Noise]

FEMALE VOICE: The simple, obnoxious piece of music that you played.

MR. PESETSKY: Yeah, I wrote it.

FEMALE VOICE: Yeah, I figured. Showing that there are differences in the importance between these notes, but if you keep--that is, if you keep the heads and drop the non-heads, you get something that not only continues to sound like a piece of music, I think I'm quoting you here, but sounds like the same piece of music.

MR. PESETSKY: The simplified version of the piece.

FEMALE VOICE: Yes, so I was thinking about what happens linguistically, if you keep the heads, okay.

MR. PESETSKY: Yeah.

FEMALE VOICE: Notice, this is exactly the opposite of what children do when they start to talk, or at least, so it seems. I mean, this is what I would ask you to comment upon.

It looks as if what the little kids are doing is dropping all the heads and keeping all of the non-heads. And that not only sounds like a piece of language, it sounds like the same piece of language. People intuitively called that, what, baby talk or various names for that sort of thing.

I can see where you get something out of keeping the heads. But you would think about that, if you were a poet, something about the structure of lines, so you get something like Jabberwocky.

That didn't seem to me like, the first guess you would make, if language and music are the same thing. That there's some reason to keep the non-heads, if you want the same piece of language.
MR. PESETSKY: If I understand the question, it has several parts to it. You didn't quite put it this way, but one would be, is there a linguistic analogue to the performance of a piece of music meeting some of the non-heads? And if not, why not?

And then the second part would be, we know that children, at a certain stage, speaking loosely, leave some words out, but they're not the words we would expect to leave out, if what they're doing is performing a simplified version of a sentence that they actually attempt to say?

On the first, more general, question, before we get to the children, I mean, if we were to try to do the trick for language, in which, let's say, let's imagine that the hypothesis, that the determiner, as the head of a nominal phrase, is correct - leave out the non-head of a DP, and just pronounce the D.

Well, let's imagine that you leave out the non-head of a T-bar or a tense, plus a verb phrase, and just pronounce the T, there are some things you have to get over in language, that you don't have to get over in music.

Two of them that are related, one is semantics, what the thing is going to mean, and the other is the lexical requirements of the heads that are left behind, and whether they license an L element or not.

I mean, if you control for that, and I say to you--I ask you a question like, did the person standing in the corner read the book that I gave him yesterday, and your answer is, he did, there is an analysis of that, in which he is a D, resuming the person who is standing in the corner, in the first phrase, and did is a spelling out of a T, that would otherwise do morphological stuff with the verb.

But I had to switch from the, as in, the person standing in the corner, to he, and I had to pronounce did, which we would not otherwise pronounce, due to lexical requirements of the D's and the T's in question. But it's not out of the question to suppose that that kind of production, and maybe even some instances of anaphora, in general, are exactly that for language.

I've never thought of this before, so maybe the idea will break down with the next question. But it's possible that there are examples of that sort, that one can think about.

I mean, as to children, well, you're the expert and I'm not, but my understanding is that, what I teach in Ling 101 or something, is that what's being typically left out is something like the function words versus the lexical categories or the phonological clitics in some theories, and not the full
forms.

And again, the functional, lexical distinction is one that refers to how we've learned lexical items, which is something that doesn't happen in music. And even the clitic, non-clitic distinction, putting some issues aside, is something, again, that's lexical and not part of music.

It could be that we wouldn't expect an--whatever it is that children are doing, and there are a bunch of theories, we wouldn't expect to find an analogue of that in music because they're not really leaving out non-heads or anything; they're leaving out morphemes that they memorized, of a certain type, but there's nothing like that in music at all.

You could ask a different question, which is, are there productions, aside from an anaphora, that maybe are characteristic of children, that mirror the variation stuff that I was talking about? But then if the answer is, no, you would have to ask why we would expect that to be?

Thank you. Yes?

MALE VOICE: A related question.

MR. PESETSKY: Yeah.

MALE VOICE: --. The general theory of Lerdahl and Jackendoff was initially or has been accepted as a theory of really listening or cognition, cognitive analysis--

MR. PESETSKY: [Interposing] That's actually how they--

MALE VOICE: Right.

MR. PESETSKY: --pitched it.

MALE VOICE: And that seems to be its strength. And so I wonder then, compared to a theory of generative grammar, which gets both that analysis and production, how do you--do you see that as an issue for this proposal, that GTTM can't necessarily handle musical production? It's more about the listener?

MR. PESETSKY: The question is whether there's an issue that they all quit, and I really don't care. Is it an issue for the GTTM model, that it's not a production model?

MALE VOICE: Right.
MR. PESETSKY: The work that I was trying to align GTTM with is also not a production model. There are at least three different questions, which obviously are connected, or rather, the answers obviously should be connected.

But in the linguistic domain, there's still a lot that we don't know, and massive debates about the nature of the connection.

One is what are people doing when they perceive and understand linguistic, understand speech? Another is, what are they doing when they produce it? And a third is, what's the knowledge system, the abstract system of rules and - - ?

I may not live to see the day when these questions receive answers that connect with each other properly. There are plenty of proposals out there, and I'm not an expert on them.

Over on the music side, I think there's a lot more work on aspects of perception. There's a hugely growing field in cognitive psychology and neuroscience and music, that's looking at details of the on-line and real time perception of the musical surface.

I guess there are some studies of production. I remember articles by Bruno Repp [phonetic] about how skilled pianists do and do not do things.

But I think the discussion of the rule system should be the domain of the field called music theory, and in some sense, it is, but not necessarily in the sense recognizable to generative grammar.

Our goals, though it may seem big and bold, and identity thesis and all of that, were actually modest and practical. We weren't, for example, what linguists should be doing, but simply saying, look, we know, I mean, this is our area of expertise, a fair amount about the rule systems underlying the structures of language.

But let's go back to GTTM and see if we can use their insights, though pitched as a theory of perception, and recast them as telling us something about the rule system, and then step back and see what's similar and what's different.

That's actually all we've done. But the question that you're asking sort of names a number of very important tasks for the future, that it's not going to be me with my particular set of expertise, that's going to accomplish those tasks, but that probably are relevant for both music and for the language
side as well. So that's a very good question.

In the back, and then Mark in the front. Yeah?

MALE VOICE 1: Do you have any thoughts about singing?

MR. PESETSKY: Do I have any thoughts about singing? Only that my friends and family don't want me to do it.

[Laughter]

MR. PESETSKY: What do you have in mind?

MALE VOICE 1: I don't know. It just seems like, singing seems like some kind of synthesis between the two sides--

MR. PESETSKY: [Interposing] Oh, I see. You mean, singing with words, with lyrics?

MALE VOICE 1: Yeah, like, songs.

MR. PESETSKY: All right. One thing that I've said nothing at all about, mainly because it seems like there are people doing excellent work on this already, but is the relation between musical meter and poetic meter, and then what's called, the text setting problem.

That has had a lot of work, a lot of really interesting work on it, trying to figure out, you know, given that there are deep similarities of some sort, or there's potential similarities between metrical use of language and meter and music, what people do, real time, when they have a hunk of language text and they have to match it to some pre-given metrical format.

The example that I hear--and I don't know if you've heard this from John Haley. I don't know if it's original with him or it goes back to Mark or to Bruce Hayes or to somebody else.

But what do you do with a drunken sailor? Which has multiple verses with very different numbers of syllables. What do you do with a drunken sailor? What do you do with a drunken sailor? And notice, I'm not singing. To, hey, ho, up he rises, which is significantly fewer syllables, yet somehow match to this.

That's actually an area where I haven't exactly seen the identity thesis stated the way we stated it, but where there are very obvious connections of some sort between language and music. And that is a topic of sort of active and
well-developed investigation.

If you're interested in that, you know, if you Google the text setting problem, you'll get to the relevant papers by Bruce Hayes, a very good phonologist at UCLA, and by other people that will, I think, address your question much better than I could. Thank you. Mark, there's a microphone coming your way.

MARK: I think you were a bit unfair to Schenker and to traditional music theory in saying that they don't address the question of adjacency of dominant to tonic, or given the explanation for it.

Certainly, what I learned many years ago in music theory was that it had to do with voice leading, and in particular, with the leading tone in dominant wanting to be adjacent, to establish the tonic note.

That might be wrong, but I think it's false that they don't consider the question.

MR. PESETSKY: The question is, if in traditional music theory, it's claimed that voice leading puts--see, I'm actually not understanding the observation.

It will be true that voice leading will constrain the chords that like to be next to each other. Why does it follow that, to have the, whatever the Rameau quote said about the feeling of conclusiveness, that the chord that precedes the tonic has to be, in fact, the dominant?

Let me say this in a better way. In an average piece of music, including the Bach Prelude that I displayed, there's more than one occurrence, say, of the C major chord of the tonic, but not always preceded by the dominant.

What is it about a total piece that privileges, let's say, all the rules of voice leading are satisfied, but that privileges or demands that dominant immediately precede the tonic at the end of the piece? And it can't be voice-

MARK: [Interposing] It contains the seven scale tone, which establishes the tonic. I mean, that could be a wrong theory, but it's a theory.

MR. PESETSKY: But if you put it that way, it's restating the problem. That is, why could it not establish the tonic in a Schenkerian world, by heading a larger musical phrase that happens to be the--

MARK: [Interposing] Because voice leading is about adjacency.
MR. PESETSKY: We'll talk afterwards.

[Laughter]

MALE VOICE 2: At this juncture, I'd like to invite everybody to the reception and the discussion can continue over - - .

MR. PESETSKY: Thank you very much.

[Applause]

[END RECORDING]