The Recursive Syntax and Prosody of Tonal Music

Jonah Katz & David Pesetsky (MIT)

1. A consensus about music and language

Two conflicting intuitions...

A. Music and language are distinct cognitive domains.
Language does not make use of octave-based pitch-collections (scales); music lacks truth-conditional semantics; everyone can talk but not everyone can carry a tune; etc. (Jackendoff 2009)

B. Music and language are deeply similar.
Both are rhythmically organized in commensurable ways (cf. metered poetry); both are law-governed complex systems; both show variation across cultures; both have an emotive-expressive function; etc.

...correlate with two traditions in the discussion of language and music

1. **two distinct academic disciplines studying the formal structure of music and language (*music theory* and *linguistic theory*), differing in descriptive apparatus and intellectual goals.**

2. **a speculative, informal literature linking music to language.**

> "From that time to this, the notion of a worldwide, inborn musical grammar has haunted me; but I would never have dreamed of basing a lecture series on a notion so ill-defined, and apparently nonverifiable, if it were not for the extraordinary materials that have appeared in recent years on the similar idea of a universal grammar underlying human speech. I have been profoundly impressed an encouraged by the burgeoning of thought in this relatively new linguistic area — an area that might be called Chomskian...This philosophical science called linguistics seems to have become our newest key to self-discovery. With each passing year it seems to substantiate ever more convincingly the hypothesis of innate grammatical competence (as Chomsky calls it), a genetically endowed language faculty which is *universal*. It is a human endowment; it proclaims the unique power of the human spirit. *Well so does music.*"
>
> (Leonard Bernstein, *The Unanswered Question: Six Talks at Harvard*)

Consensus*: *Music and language are indeed distinct cognitive domains.

Consequence for Cognitive Science

Because the cognitive sciences have inherited the consensus view that the formal systems governing linguistic and musical structures are distinct...

- ...most research that compares language and music concerns issues of *use* rather than formal structure of rule systems: e.g. Are memory resources shared by the processing of music and language? Is there selective impairment of musical and linguistic skills? etc.
- ...and the occasional discovery that music and language do share some resource or property *P* is taken as a contribution to the topic of *domain-specificity*, and a possible indication that *P* is domain-general in some fashion (Patel 2003, Fedorenko et al. 2009, among others.)

---

2. Against the consensus: the *Identity Thesis for Language and Music*

Central claim of this talk:

- When the formal structures underlying language and music are properly compared, the consensus view turns out to be wrong. **The "syntactic components" of music and language are formally *identical.***

Our starting point:

- Lerdahl, Fred, and Ray Jackendoff. 1983. *A generative theory of tonal music*. (**GTTM**) — still the most complete and insightful formal account of tonal musical structure.

What is the same about music and language:

- **Structure:** When the GTTM model is reworked in a format commensurate with linguistic theory and compatible with its goals, we see that the syntactic structures of both music and language are the result of **iterating, binary recursive Merge**, with the additional property of **endocentricity (headedness)**.

- **Information-flow and Architecture:** The structures produced by Merge interact with the computations of other components of the grammar similarly in both music and language.

What is not the same:

- **Building blocks:** The building blocks of linguistic syntax are **lexical items** (arbitrary pairings of sound and meaning) — which have no obvious counterpart in music. The building blocks of tonal musical structure concern **pitch-class and chord quality** — which likewise have no linguistic analogues.

In a nutshell:

- **We will suggest that what language and music have in common is not their building blocks, but *what they do with them*.**

(1) **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 language; pitch-classes and pitch-class combinations in the case of music). In all other respects, language and music are identical.

***Other ways in which music resembles language**:*
Universality across the species; *acquisition profile* (e.g. early preference for consonant intervals; Zentner & Kagan (1998), Trainor, Tsang & Cheung (2002)); *species-specificity* (Hauser & McDermott (2003)), hip-hop parrots notwithstanding. See Patel (2008).

Our conclusion: Music and language count as *the same cognitive domain* with respect to many of the questions considered at this conference.

Internal Merge?
- The thesis that music and language constitute the same cognitive domain — both involving Merge — leads us to ask whether the two types of Merge found in language are both present in music: ***Internal Merge*** (syntactic movement) as well as ***External Merge*** (basic phrase structure). We will argue that they are — in particular, that...

The phenomenon of ***cadence*** exemplifies Internal Merge (in particular *head movement*) in the musical domain.

But didn't Lerdahl and Jackendoff reject the Identity Thesis in GTTM?
"[T]he generative music theory developed here does not look much like generative linguistics." (GTTM 307)

Our responses:

1. **In several technical respects, GTTM was a *decade ahead* of generative linguistics:** Chomsky's "Bare Phrase Structure" (Chomsky 1995a) and Kayne's "Connectedness" (Kayne 1983), which inform current ideas about Merge, all lay in the future. Several aspects of GTTM more closely resemble current ideas in generative linguistics than they resemble work of the late 1970s (when GTTM was developed).

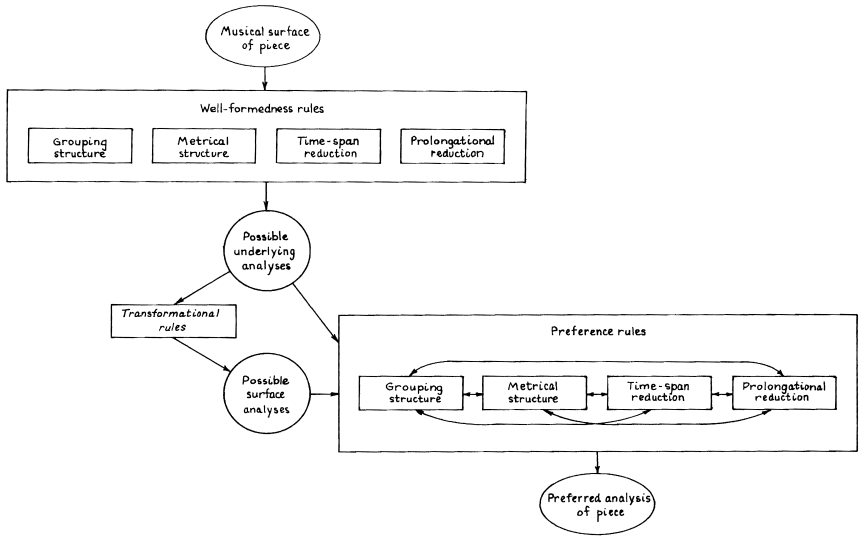
2. **Conceptually, the goals of GTTM differ from the goals of most generative linguistics.**
   This difference affected how the model was presented and which aspects were emphasized. GTTM was explicitly concerned with the following question: "Given a piece of music, what analysis does a skilled listener assign to it?"

   It was less concerned with the question: "What is the logical structure of the *rule systems* that license the analyses provided by listeners for pieces?" — which would be a question more analogous to those at the heart of most generative linguistics.

So is this talk just a book report on GTTM?
- Not really: part of our project has been a thoroughgoing ***realignment*** of the GTTM model with generative linguistics, which allows both similarities and differences to be examined more meaningfully than GTTM's presentation allows. *If we are correct, this realignment supports the Identity Thesis.*
- This realignment, in turn, allows us to extend the GTTM model to offer our ***Internal Merge account of cadence***, with which we will end the talk. *If we are correct, this too supports the Identity Thesis.*

3. Prolongational Reduction: Musical Merge

(2) **The "Official Architecture" of the GTTM model**



Though (2) does not look like models for language, similarities emerge if we focus on the actual function and interaction in practice of specific components.

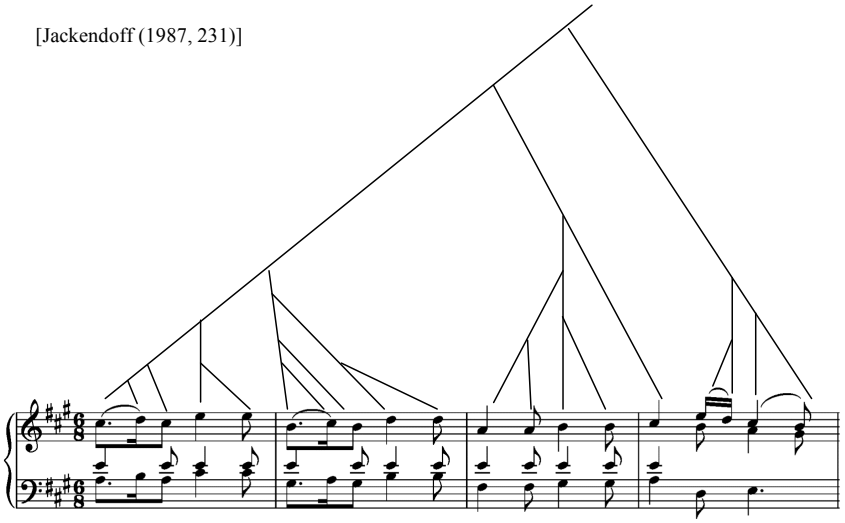
Prolongational Reduction

A hierarchical representation of tension/relaxation
- The harmony of a piece of music is organized into a **binary-branching headed tree** called ***Prolongational Reduction* (PR)**, whose heads are chosen so as to minimize distance between sisters. *Minimize* distance = *maximize* stability.

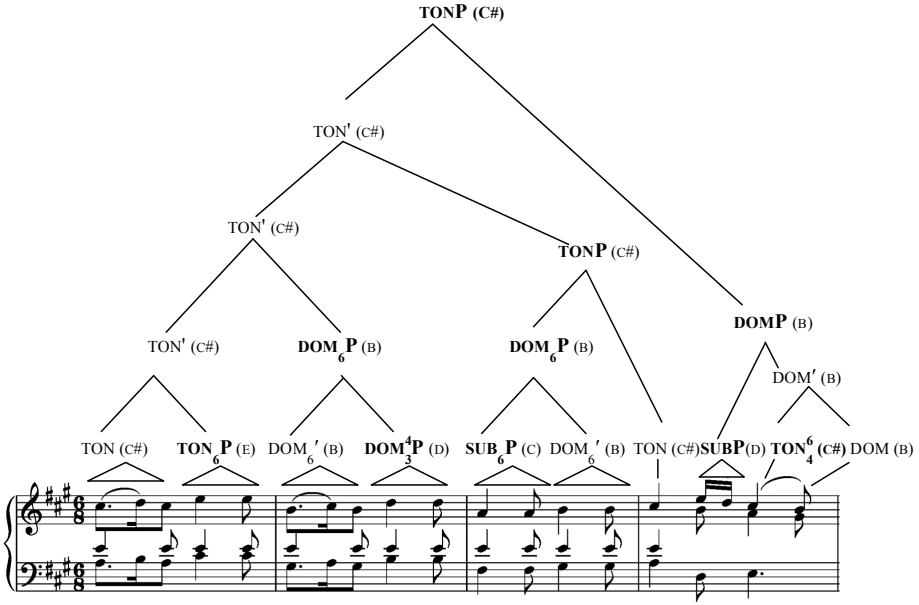
Relevant notions
- **Distance among chords:** The distance between two chords is measured in a somewhat complex *pitch-space* that takes account of the notion *key* (Krumhansl 1990; Lerdahl 2001).
- **Relative stability:** The relative stability of connections between chords in a given key is determined by their distance from each other.

(3) **PR structure for Mozart piano sonata K. 331: GTTM notation**

[Jackendoff (1987, 231)]



(4) **PR structure for Mozart piano sonata K. 331: linguists' notation**[1]



[1] TON=*tonic* (chord built on first note of scale), DOM-*dominant* (built on fifth note of scale); SUB=*subdominant* (here: built on fourth note of scale). See Katz & Pesetsky (in prep.) for an argument that it is actually the pitches of a chord that project syntactically, not these chord labels — but this distinction is not relevant for the point made here.

The intuitions modeled by PR

- PR models patterns of harmonic tension and relaxation on the hypothesis that for any pair of sisters (α, β), the more tense member of the pair is the one that is harmonically more distant from the chord that minimally c-commands α and β.
- These patterns reflect hierarchy (Lerdahl 2004) in that...
  - ...the tension associated with an event depends in part upon its structural context in a PR tree.
  - ...these patterns are not explicable in terms of 'flat' linearity or string-adjacency.
- Predictions of PR confirmed experimentally by Lerdahl & Krumhansl (2007).

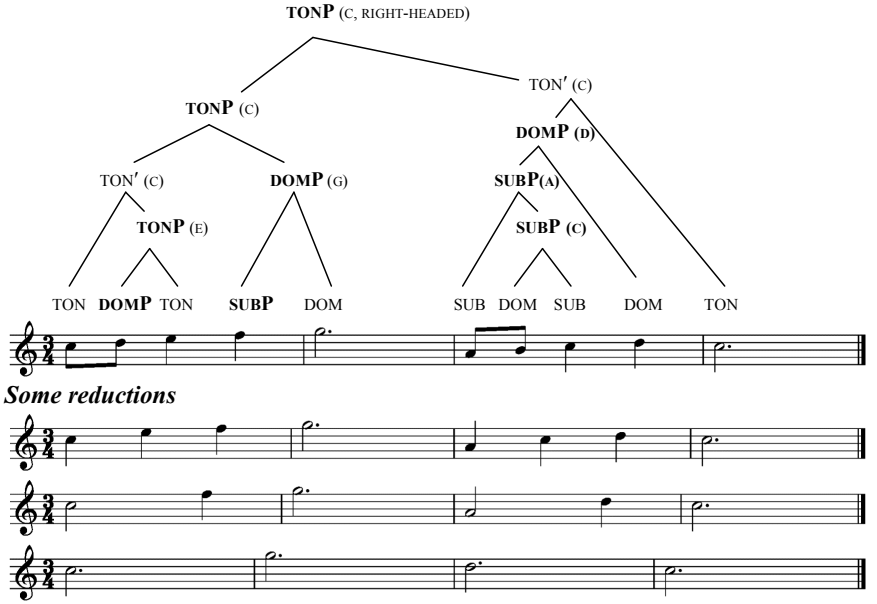
The formal structure of PR

- **Trivially, these structures can be viewed as arising from *recursive Merge.***
  - ***Merge*** because it involves hierarchical structure.
  - ***Recursive*** because the rules governing sisterhood and choice of head appear to be the same at all levels of structure (cf. Patel 2008, 264-5).

Simple Demonstration of Recursive Property

- ***Reduction:*** Because the organizing principles of PR are identical at every level, a performance that substitutes the head of any constituent for the full set of terminals of that constituent (a ***reduction***) will still "sound like music". This is the principle behind "***theme and variations***" (with a caveat discussed in the next section).

(5) **The sound of recursive syntax...**



Some reductions

***Non-local interactions among non-adjacent elements*:** The fact that reductions "sound like music" illustrates the possibility of **non-local interaction** that follows from a hierarchical structure that is also **headed**.

- This observation originates with the the music theorist **Heinrich Schenker** (1868-1935), whose work greatly influenced GTTM — though he stopped short of fully understanding what he had discovered (i.e. recursion in music):

(6) **Schenker's generalization**
Two chords interact harmonically (forming a tension/relaxation pair) so long as they project phrases that are sisters. The heads themselves need not be linearly adjacent.

- See Giblin (2008, esp. chapter 3), for much further discussion and arguments for recursive structures in tonal music.

But PR is not sufficient to account for all intuitions about reduction...

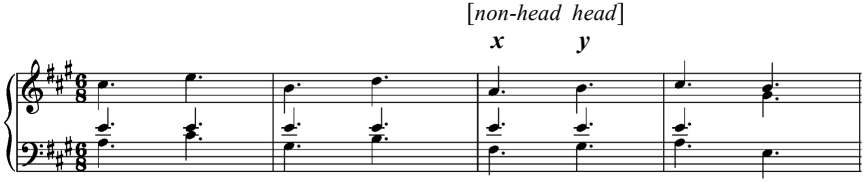
4. Musical Prosody: Time-span reduction (TSR)

Rhythm and parallelism* (musical prosody!) *also enter into reduction judgments

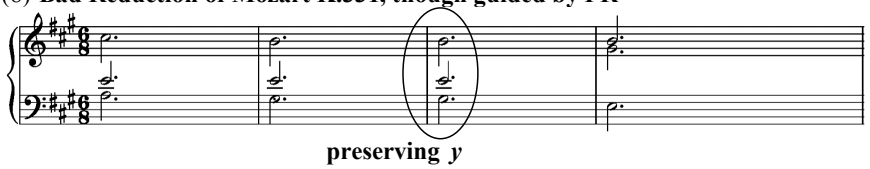
- ***An example from GTTM* (p. 164)*:***
On the one hand: The E-major (DOM$_6$) chord labeled *y* below is chosen by PR principles as the head of the bracketed constituent, rather than the preceding chord *x* (SUB). This accords with the general fact that DOM chords (built on the 5th degree of the scale) strongly select for a following TON.

But: A reduction that retains *y* and loses *x* is clearly a less successful reduction than a reduction that retains *x* and loses *y*. Clearly relevant is the fact that the *x* occupies a ***metrically stronger*** position than y, and also parallelism with the metrically strong position of the chords retained in previous bars:

(7) **A first reduction of Mozart K. 331 (ok!)**

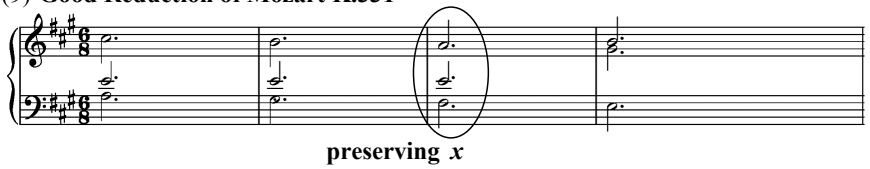


(8) **Bad Reduction of Mozart K.331, though guided by PR**



preserving *y*

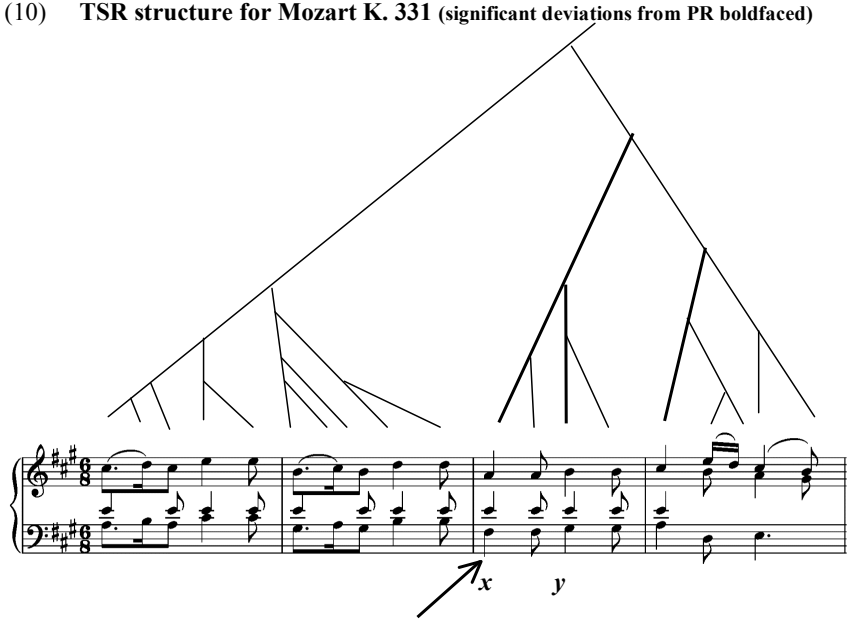
(9) **Good Reduction of Mozart K.331**



preserving *x*

GTTM's Solution:
an independent representation of prominence, Time-Span Reduction (TSR)
— also a hierarchically organized, headed tree.

(10) **TSR structure for Mozart K. 331 (significant deviations from PR boldfaced)**



the chord that needs to be retained, even though it's a PR dependent on the chord to its right

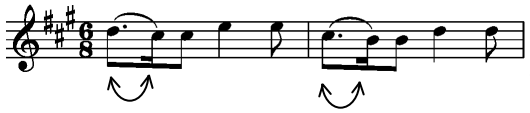
- **Metrical strength, spacing among events, and parallelism all play a role in building TSR...**

- **...but PR's representation of harmony also plays a major role in constraining TSR**.

  Demonstration: ***If the first and second chords of bars 1 and 2 of Mozart K. 331 are reversed, our intuition about reduction remains unchanged*** — despite the fact that it is a weak beat that is being retained.

  Properties of harmony (e.g. pieces start with a tonic) take precedence:

(11) **Reversed version of Mozart K. 331 (weak/strong beats exchanged)**



Relation of PR to TSR

- ***How TSR represents prominence:*** The prominence of a chord/pitch α in TSR corresponds to the ***number of nodes that separate the maximal projection of α from the root of the tree***.

  We call this the ***Root Distance* of α, RD(α):**

(12) **Root Distance number (RD)**
RD(α) = the number of nodes that dominate the maximal projection of α.

(13) **Prominence**
Lower RD number corresponds to greater prominence.

- ***How PR constrains TSR:*** GTTM identifies a significant constraint that PR imposes on TSR (though GTTM phrases the constraint quite differently):

(14) **Region condition**
In the string (...α...*e*...β...), if α and β are the heads of sisters in PR, RD(*e*)≥RD(α) & RD(*e*)≥ RD(β).[2]

[2] A significant exception to the Region Condition is allowed when two sister nodes in PR are formed by identical chords (called a *strong prolongation*; this exception corresponds to the *Interaction Principle* of GTTM). The full statement of (14) should thus include:

...unless RD(*e*)-RD(α)=1 or RD(*e*)-RD(β)=1; and *e* strongly prolongs α.

In plainer terms: Given the binarity of branching in PR, *e* in (14) must be dominated by αP or βP (or both). It may not be promoted in TSR prominence past either one of them.

Function of the Region condition: Plausibly facilitates listener's extraction of PR constituency from the signal, by guaranteeing that if α and β are interrupted by a more prominent *e*, then α and β may not be the heads of sisters.

In (9) (Mozart K.331), the promotion of x over y in TSR is allowed precisely because x did not linearly intervene between y and the sister of yP.

- ***Two working hypotheses about architecture of the grammar*** **(take your pick)*:***

  Derivational: PR→TSR, or
  Representational: (PR, TSR) co-satisfies a set of constraints, including the Region Condition

- ***Conclusion:*** There is a systematic relationship between ***musical prosodic prominence*** **(TSR)** and the ***syntactic relations given by Merge*** (**PR**).

---

5. Linguistic Prosody

Observation: **TSR is formally similar or identical to representation of linguistic prosody (GTTM, Ch. 12) — and represents similar notions of prominence and phrasing.**

Our proposal: **The relation between PR and TSR *is* the musical counterpart to the relation between linguistic syntax and prosody.**

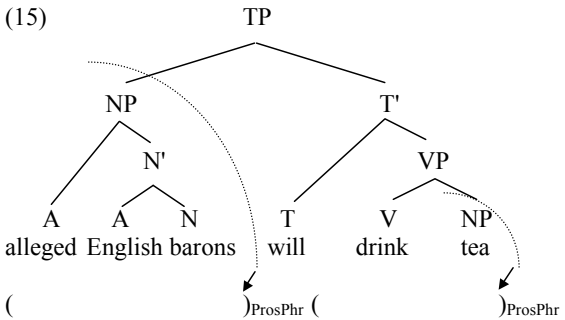
- ***How prosodic structure represents prominence:*** Prosodic structures in language are hierarchical and headed, but not identical to corresponding syntactic structures.

- ***Two working hypotheses about architecture of the grammar*** **(again, take your pick):**
  Derivational: syntax→prosodic structure (Nespor & Vogel 1986, Selkirk 1984, Hayes 1989), or
  Representational: (syntax, prosodic structure) co-satisfies a set of constraints, (Selkirk 2000, Truckenbrodt 1999)

- ***How syntactic structure constrains prosodic structure*** **(Selkirk, *passim*):**
  Edges of certain prosodic categories are aligned with the edges of certain syntactic categories.

  Phases: In English, the ***right edge*** of each NP [or more generally, all syntactic *phases*] that bears stress is aligned with a ***prosodic phrase boundary***.

(15)



Prosodic word: In English the right edge of each minimal unit composed of a word and any clitics attached to that word is aligned with a ***prosodic word boundary***.

Prosodic headedness (greatest prominence in category):
Determined by phonological or syntactic/lexical properties of the constituents.

Word and prosodic word: In English, word-level stress is determined by a distinct phonological system.

Above the word: Head prominence in English is generally assigned to the ***rightmost stressed syllable in each constituent*** (with important exceptions). We call this syllable the ***phase-marker*.**

6. Musical PR-TSR mapping = Linguistic syntax-prosody mapping

- ***A similarity between music and language:***
  Much as the relation in music between TSR and PR requires a decrease in TSR prominence *between heads of PR-sisters*, so the relation in language between syntax and prosodic structure requires a decrease in prosodic prominence *between each phase-marker and the next phase-marker,* and *between each non-clitic and the next non-clitic*.

- ***Thus, there already is a region condition for language just as there is for music:***
  Both domains strictly limit the ability of an event *e* in "...α...*e*...*β*..." (where α and β are the heads of sisters) to be promoted in prominence over α and β:

(16) **Prosodic Prominence for music and language**
α exceeds β in prosodic prominence iff RD(α) < RD(β).

(17) **Region conditions for music and language**
For every pair of distinct PR/syntax events (α, β), such that:

a. ***Music*:** α and β are the heads of sisters; or

b. ***Language*:** α and β are phase-markers, and β is the phase-marker linearly closest to α; or

c. ***Language*:** α and β are non-clitics, and β is the non-clitic linearly closest to α

if an event *e* linearly intervenes between α and β, both α and β exceed *e* in prominence.

- ***A difference between music and language:***
  Music: The region within which the musical condition requires a decrease in TSR-prominence is defined in terms of PR *sisterhood*.

  Language: The region within which the linguistic conditions require a decrease in TSR-prominence is not dependent on syntactic sisterhood, but on the *lexical properties of particular syntactic items*. This dependence is evident in the case of clitics/non-clitics (a lexical property) — but the lexicon also dictates which lexical items project *phases* (from which an element's status as a phase-marker is derived).

- ***Is this difference a problem for the Identity Thesis?***
  1. The function of Region Conditions in both domains is to constrain non-isomorphy between prosodic/TSR structure and syntactic/PR structure — and perhaps to encode some information about major phrase boundaries using prominence distinctions as a tool. **So Region Conditions serve similar purposes in language and in music.**

  2. Since music lacks a lexicon, the region conditions applicable in language could not in principle be implemented in music as well. As a consequence: **formally, conditions (17b-c) might apply in music too**, **but vacuously** — because no chord is lexically a phase-head and no chord is lexically a clitic or non-clitic.[3]

  3. This leaves (17a) to worry about. (Work in progress.)

[3] RD number for language appears to count only designated prosodic phrase nodes, while musical RD appears to count every node. In Katz & Pesetsky (in prep.) we suggest that this too might reflect the core distinction between music and language: the absence of the notion *lexical item* in music.

7. Interim summary: comparing music with language

Building blocks are different:
- Lexical items in language, chords/pitches in music.

Syntactic component (PR, linguistic syntax) is the same:
- Both language and music combine elements with *iterated, binary, recursive Merge* — plus *headedness*.

What's it all about?
- ***PR sisterhood in music*** reflects harmonic tension/relaxation.
- ***Syntactic sisterhood in language*** reflects satisfaction of lexical properties.

Prosodic component (TSR/linguistic prosody) and its relation to the syntactic component is similar, perhaps identical:
- Constrained by *Region Condition*, which imposes a degree of isomorphy with syntactic structure, while allowing some deviation.

What's it all about?
- ***RD number for TSR in music*** is related to harmonic structure and cued by perceptual/production factors like metrical strength, grouping and parallelism.
- ***RD number for prosody in language*** related to syntactic structure, and cued in similar ways.

---

8. Internal Merge and the notion *Cadence*

Given straightforward expectations concerning syntactic constituency (e.g. *V merges with its complement argument*)...

...and given straightforward ideas about how syntactic structures should be linearized (e.g. *when two elements merge, the right edge of one element aligns with the left edge of the other*)...

<u>Language</u>
- ...it is interesting that in many cases the otherwise predictable relation between hierarchical syntactic structure and phonological surface breaks down:

(18) **French V separated from complement by negation**

la fille	n' achèter-a	[$_{VP}$ **pas**	___	le livre].
the girl	*prt* buy-FUT-3SG	**not**		the book

'The girl will not buy the book.' (Emonds 1978, Pollock 1989)

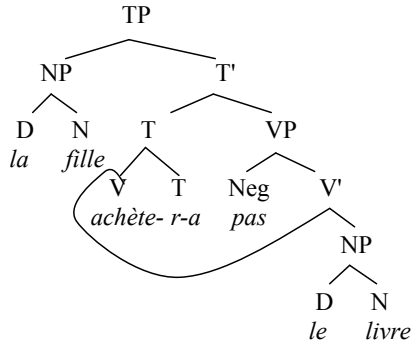
- **Hallmarks of the phenomenon**

1. Finite V is pronounced adjacent to (and tightly coupled with) Tense, yet...
2. ...has a normal set of syntactic dependents, linearized where we also expected V to be linearized. Thus, V apparently does head its own VP.

- **Explanation: *Internal Merge* of V to T (here, "head movement")**
  In an example like (18), the verb *acheter* 'buy' ***did merge with its complement*** *le livre*, as expected by the normal laws that link thematic role assignment to syntactic position...

  ...and then ***merged a second time —*** with T (located outside VP):

(19)



- **Other hallmarks of the phenomenon**

3. **Obligatoriness of the movement**
   Movement of finite V to T in French satisfies some *need* of an element in this structure — probably T, since non-finite verbs do not obligatorily raise to T.

   Thus, the T-V relationship involves some *alteration* in the features of T.

   [***How it works***: An Agree relation (Chomsky 2005b, 2000) between an unvalued tense feature of T and its valued counterpart on V assigns a value to the feature on T. As a language-particular consequence of this instance of Agree (*EPP*), V must Internally Merge with the T that it has just agreed with (Pesetsky & Torrego 2007).]

4. **Tight coupling of V with its host, which function as a phonological word**
   A general characteristic of head-to-head movement (much discussed, cf. Matushansky 2006). This property may be related to the property of T that motivates the movement in the first place (the "stray affix filter" of Lasnik 1981).

Music

- Are there cases in music in which an otherwise expected relation between PR hierarchy and linear order breaks down?

- **What we might look for if music shows *head-to-head Internal Merge* like (19):**
  1. Some chord X must be performed **adjacent** to a chord Y, yet...
  2. ...X has a **normal set of syntactic dependents** of its own, linearized normally — and thus apparently also **heads its own phrase** (an XP).
  3. **Obligatoriness**, related to some necessary ***alteration*** in the features of Y, required in order for the derivation to succeed; and
  4. **Tight coupling of X with its host Y**, such that they function as an indivisible unit for other purposes (cf. the notion *word*)

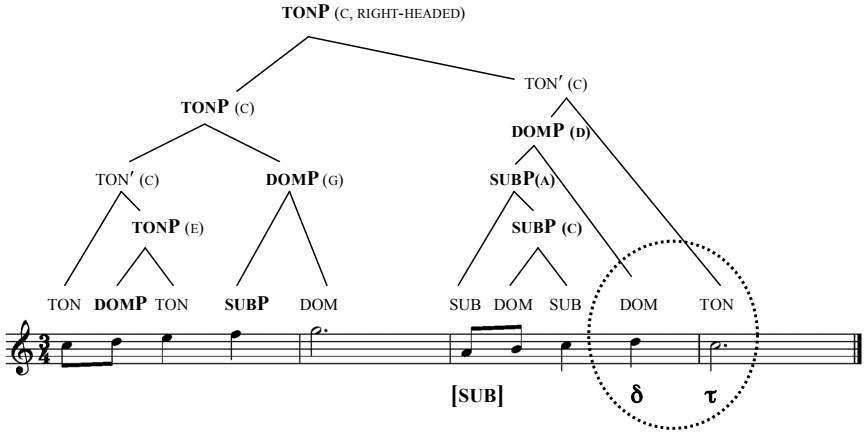
- ***Cadence* ("perfect cadence")**
  A sequence of two chords (δ, τ), where τ is a chord built on TON (first note of the scale) and δ is a major triad built on DOM**,** the pitch five scale-steps higher than τ.

> The cadence is a major structural signpost that establishes key and marks the ends of sections of musical pieces.

Crucially, a cadence has the following additional properties:

1. δ must be **adjacent** to τ , yet...
2. ...δ has a **normal set of syntactic dependents**, linearized normally — and thus apparently also **heads its own phrase** (δP). The presence of at least one dependent of δ — a **subdominant [SUB]** (built most often on scale-degree 4 or 2) — is exceedingly common.

(20) **The melody from (5) and its PR structure (cadence circled)**



3. The participation of τ in a cadence (δ, τ) is a precondition for **establishing the key as "the key of τ"**.

   How this works: **τ is marked [+TON] by δ**. This feature plays a role in the structural condition on key-establishment in (21):

(21) **Key establishment**[4]
If a node in the key of τ, it is dominated by a phrase headed by a [+TON] instance of τ.

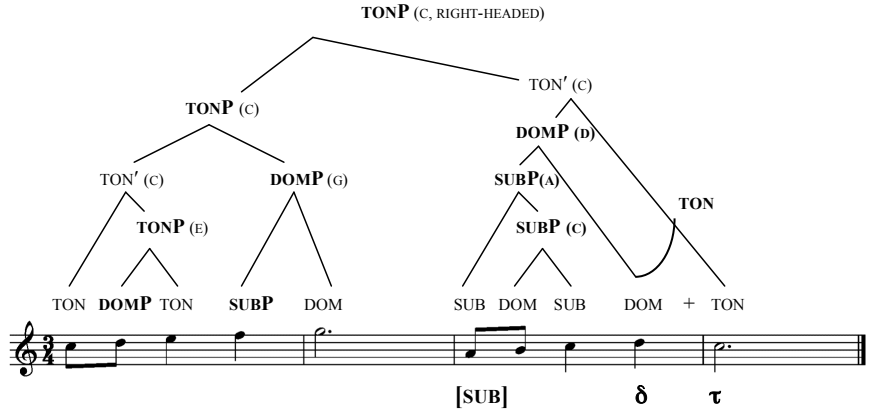
Corollary: A tonal piece must be in some key, so a piece must be headed by a [+TON]-marked chord. *Western tonal pieces are right-headed — consequently a piece in that idiom must end with a cadence.*

- **Why is there a head-adjacency condition on [+TON]-marking?**
  Why is it not sufficient for a phrase *headed* by δ to Merge with a phrase *headed* by τ? All other musical relations between heads appear to require only adjacency between their projections — not the heads themselves.

- **Solution: *Internal Merge of δ with τ required for tonic-marking***
  In an example like (5)**, the head of the final DOM phrase (δ) first Externally Merged with SUB (the chord that precedes it)...**

  ...and then the head δ *merged a second time* — with the head τ:

(22) **Cadential Internal Merge of δ to τ (Head-Movement)**



[4] We state (21) as a conditional, rather than a biconditional, because PR subtrees may be in a different key from a larger structure that contains them (a phenomenon called *tonicization* or *modulation*). It is in fact the smallest domain satisfying (21) that establishes what key a given node is in. These issues are discussed at length in Katz & Pesetsky (in prep.), where they are argued to justify the localization of the notion *key* in an interpretive component (cf. LF for language), rather than in PR itself.

- ***Key point***:
  Even if the DOM phrase headed by δ in (22) were *not* head-final, the result would sound the same, because δ would move rightward to τ.

  [***How it works***: Agree between δ and τ marks τ as [+TON]. As a consequence of this instance of Agree (*EPP*), δ must Internally Merge with the τ that it has just agreed with.]

On the indirectness of the argument for Internal Merge

Because there is no subcategorization or thematic-role assignment in music (and no piece-wide setting head-first/head-last setting), it is not possible to pinpoint the position where we might have expected to hear δ if it had not undergone cadential head-movement.

This contrasts with languages like French, where we can use principles of selection and head-complement ordering to localize the position of V in (19).

For this reason, the evidence for head movement in music must be more indirect than corresponding evidence in language (and there is nothing we can do about that!).

To summarize the evidence:

- the presence of an absolute string-adjacency condition on a pair of heads,
- ...linked to the valuation of a feature on the higher head,
- ...in a component of grammar where *all other interactions* between elements X and Y require only sisterhood between *constituents headed by X and Y* — not linear adjacency between X⁰ and Y⁰ themselves.

- **Tight coupling of δ with τ?**
  For purposes of the Region Condition, δ behaves as if RD(δ) = RD(τ)-1.

  Evidence from reduction*:* If a given reduction of a cadenced piece of music includes any chords other than the final tonic τ, it also includes the dominant δ — even if δ was rhythmically non-prominent and deeply embedded harmonically.

  [This recasts a GTTM stipulation about the PR-TSM mapping called ***cadential retention***.]

9. Conclusions

- ***Results:*** When we view GTTM's discoveries and proposals through the prism of modern generative linguistic theory, strong formal similarities emerge between music and language.

  The formal theory of tonal music that emerges from our recasting and extension of GTTM *does* "look much like generative linguistics" — in particular, boasting a central syntactic component that combines elements by means of **iterated, recursive Merge,** both External and Internal.

- ***The Big Picture:*** Support for the **Identity Thesis for Language and Music**, repeated below:

(1) **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 language; pitch-classes and pitch-class combinations in the case of music). In all other respects, language and music are identical.

— or at least: support for the fruitfulness of the Identity Thesis as a strategic starting point for research on musical and linguistic grammars.

- ***Some Open Questions:***
  If the same combinatory system that recursively combines lexical items in language is also put to work recursively combining musical units from tonal pitch space...

  1. Do other cognitive activities employ the same recursive combinatory system? Are lexical items and musical pitches privileged as "raw ingredients" for the system?

  2. Where do the special properties of pitches relevant to the syntax of music come from? Are pitches as different from lexical items as they appear to be?

  3. If music and language are a single cognitive domain in the sense claimed here — Is music a by-product of language — in effect, a kind of *language game*? Or is language perhaps a kind of music game...?

References

Bernstein, Leonard. 1976. *The Unanswered Question: Six talks at Harvard*. Cambridge, Mass: Harvard University Press.

Chomsky, Noam. 1995a. Bare phrase structure. In Gert Webelhuth, ed. *Government binding theory and the minimalist program*. 383-439. Oxford: Oxford University Press.

Chomsky, Noam. 1995b. *The minimalist program*. Cambridge: MIT Press.

Chomsky, Noam. 2000. Minimalist inquiries. In Roger Martin, David Michaels, and Juan Uriagereka, eds. *Step by step: Essays on minimalist syntax in honor of Howard Lasnik*. 89-156. Cambridge, Mass.: MIT Press.

Emonds, Joseph. 1978. The verbal complex v′ - v in French. *Linguistic Inquiry.* 9:151-175. [http://www.jstor.org/stable/4178050]

Fedorenko, Evelina, Aniruddh Patel, Daniel Casasanto, Jonathan Winawer, and Edward Gibson. 2009. Structural integration in language and music: Evidence for a shared system. *Memory and Cognition.* 37:1-9. [http://mc.psychonomic-journals.org/content/37/1/1.full.pdf+html]

Giblin, Iain. 2008. *Music and the generative enterprise*. Doctoral dissertation, University of New South Wales.

Hauser, Marc D., and Josh McDermott. 2003. The evolution of the music faculty: A comparative perspective. *Nature Neuroscience.* 6:663-668.

Hayes, Bruce. 1989. The prosodic hierarchy in meter. In Paul Kiparsky, and Gilbert Youmans, eds. *Phonetics and phonology, volume 1: Rhythm and meter*. 201-260. Orlando, FL: Academic Press.

Jackendoff, Ray. 1987. *Consciousness and the computational mind*. Cambridge, MA: MIT Press.

Jackendoff, Ray. 2009. Parallels and nonparallels between language and music. *Music Perception.* 26:195-204.

Kayne, Richard S. 1983. Connectedness. *Linguistic Inquiry.* 14:223–249.

Krumhansl, Carol L. 1990. *Cognitive foundations of musical pitch*. New York: Oxford University Press.

Lasnik, Howard. 1981. Restricting the theory of transformations. In Norbert Hornstein, and David Lightfoot, eds. *Explanation in linguistics*. London: Longmans.

Lerdahl, Fred. 2001. *Tonal pitch space*. Oxford: Oxford University Press.

Lerdahl, Fred, and Ray Jackendoff. 1983. *A generative theory of tonal music*. Cambridge, MA: MIT Press.

Lerdahl, Fred, and Carol Krumhansl. 2007. Modeling tonal tension. *Music Perception.* 24:329-366.

Matushansky, Ora. 2006. Head movement in linguistic theory. *Linguistic Inquiry.* 37:69–109.

Nespor, Marina, and Irene Vogel. 1986. *Prosodic phonology*. Dordrecht, Holland ; Riverton, N.J.: Foris.

Patel, Aniruddh D. 2003. Language, music, syntax and the brain. *Nature Neuroscience.* 6:674-681. [http://www.nature.com/neuro/journal/v6/n7/abs/nn1082.html]

Patel, Aniruddh D. 2008. *Music, language, and the brain*. Oxford ; New York: Oxford University Press. [http://www.loc.gov/catdir/toc/ecip0715/2007014189.html]

Pesetsky, David, and Esther Torrego. 2007. The syntax of valuation and the interpretability of features. In S. Karimi, V. Samiian, and W. Wilkins, eds. *Phrasal and clausal architecture*. Amsterdam: Benjamins. [http://web.mit.edu/linguistics/www/pesetsky/Pesetsky_Torrego_Agree_paper.pdf [written 2004]]

Pollock, Jean-Yves. 1989. Verb movement, universal grammar, and the structure of IP. *Linguistic Inquiry.* 20:365-424. [http://www.jstor.org/stable/4178634]

Selkirk, Elisabeth. 1984. *Phonology and syntax: The relation between sound and structure*. Cambridge, Mass.: MIT Press.

Selkirk, Elisabeth. 2000. The interaction of constraints on prosodic phrasing. In Merle Horne, ed. *Prosody: Theory and experiment (studies presented to gösta bruce)*. 231-261. Dordrecht: Kluwer.

Trainor, Laurel J., Christine D. Tsang, and Vivian Cheung. 2002. Preference for sensory consonance in 2- and 4-month old infants. *Music Perception.* 20:187-194. [http://www.psychology.mcmaster.ca/ljt/trainor_tsang%20et_al_2002.pdf]

Truckenbrodt, Hubert. 1999. On the relation between syntactic phrases and phonological phrases. *Linguistic Inquiry.* 30:219-255. [http://www.jstor.org/stable/4179060]

Zentner, Marcel, and Jerome Kagan. 1998. Infants' perception of consonance and dissonance in music. *Infant Behavior and Development.* 21:483-492. [http://www.unige.ch/fapse/emotion/members/zentner/pdf/article3.pdf]