# The Identity Thesis for Language and Music

Katz & Pesetsky 2011

# 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."
  - Language and music share a common syntactic component.
  - Differences are in the building blocks, including differences in how the structures are processed by distinct interpretative components.
  - Irrelevant to music is propositional, truth-conditional meaning that originates in words.
  - Irrelevant to language are notions like "key" and "tonic".

# Levels of generative description for music

- Type 1. Analysis of particular pieces.
  - *A listener who hears a sequence of sounds S in terms of a musical idiom I assigns one or more analyses to S. An analysis that is assigned to S can be discerned by a variety of judgments that a listener can render about S within I. What general laws define the class of possible analyses within I of a given piece?*
  - An explicit, predictive, type-1 theory might be called a **generative parser** for I.
- Type 2. Common properties of pieces within an idiom.
  - *A listener capable of assigning an analysis in a musical idiom I to a sequence of sounds S may identify or not identify S as a piece admitted by I. What general laws define the class of possible pieces in I?*
  - An explicit, predictive, type-2 theory is a **generative grammar** for I.

# Levels of generative description for music cont.

- Type 3. Common properties of musical idioms.
  - *What is the class of possible grammars for human musical systems?*
  - An explicit, predictive, type-3 theory of possible grammars for human musical systems might be called a **Universal Grammar** for music (**UG-M**).
- Type 4. Properties common to UG-M and other cognitive systems.
  - *Which properties of UG-M are unique to music, and which are shared with cognitive systems generally regarded as distinct? For example, does UG-M share significant properties with Universal Grammar for language (UG-L)?*
  - An explicit, predictive, type-4 theory thus **distinguishes music-specific components of Universal Grammar for music** from other aspects of UG-M.

# Differences between generative theories of language and music

- The Identity Thesis most directly reflects type-4 concerns.
- GTTM (Jackendoff & Lerdahl 1983) is a type-1 theory, whereas generative grammar theories for language are type-2 and type-3.
  - "Overall the system can be thought as taking a given musical surface as input and producing the structure that the listener hears as output" (GTTM, 11).
  - This accounts for the differences noted by Jackendoff (and Patel) between the two theories (systems).
  - For direct comparison, the two theories need to be "aligned".
  - CLAIM: When they *are* aligned the important differences disappear.

# Similarities between the two systems even before alignment

- "Metrical structure" in GTTM:
  - Formally identical to the representations of linguistic stress proposed in Liberman & Prince (1977), Hayes (1995)
    - Metrical phonology.



```
                                              X
                          X                   X
      X             X     X             X     X
    X X             X     X       X     X     X
X   X X   X   X X X       X   X   X X X X   X
```

In November, the region's weather was unusually dry

Figure 3.11  A metrical grid for a sentence of English. From Selkirk, 1984.

# Similarities between the two systems even before alignment

- "Time-span reduction" in GTTM:
  - Similar to the prosodic structures of Selkirk (1980, etc.).
  - Both are characterized by hierarchical nested phrasal domains, each headed by a unique event at every hierarchical level.
  - The property of being a head in both systems involves some notion of rhythmic prominence.
  - These structures govern the relation between syntax and phonology in both systems.
- Still, the syntactic building blocks are different:
  - Music:  pitch, key, scale, interval, chord type, cadence.
  - Language: phoneme, morpheme, word, part of speech, agreement.

# Align-head rightmost in prosodic unit (C, P, I, U), stress head
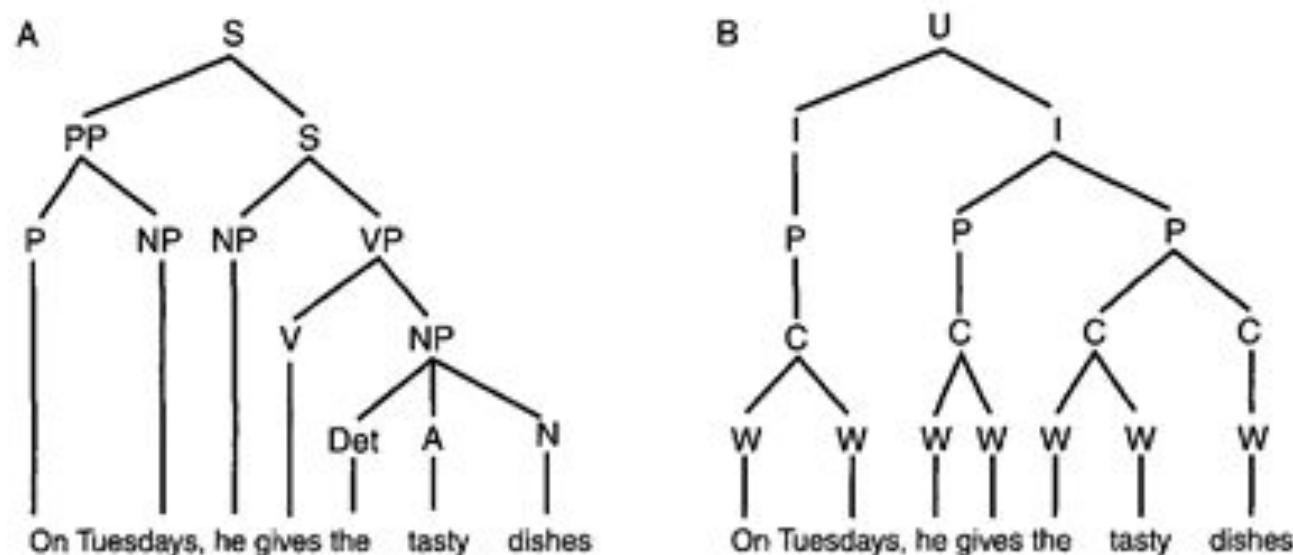
Figure 3.5 (A) Syntactic and (B) prosodic hierarchy for a sentence of English. Abbreviations for (A): S = sentence, PP = prepositional phrase, NP = noun phrase, VP = verb phrase, Det = Determiner, A = adjective, N = Noun, V = Verb. Abbreviations for (B): U = utterance, I = Intonation phrase, P = phonological phrase, C = clitic group, W = word. Adapted from Hayes, 1989.
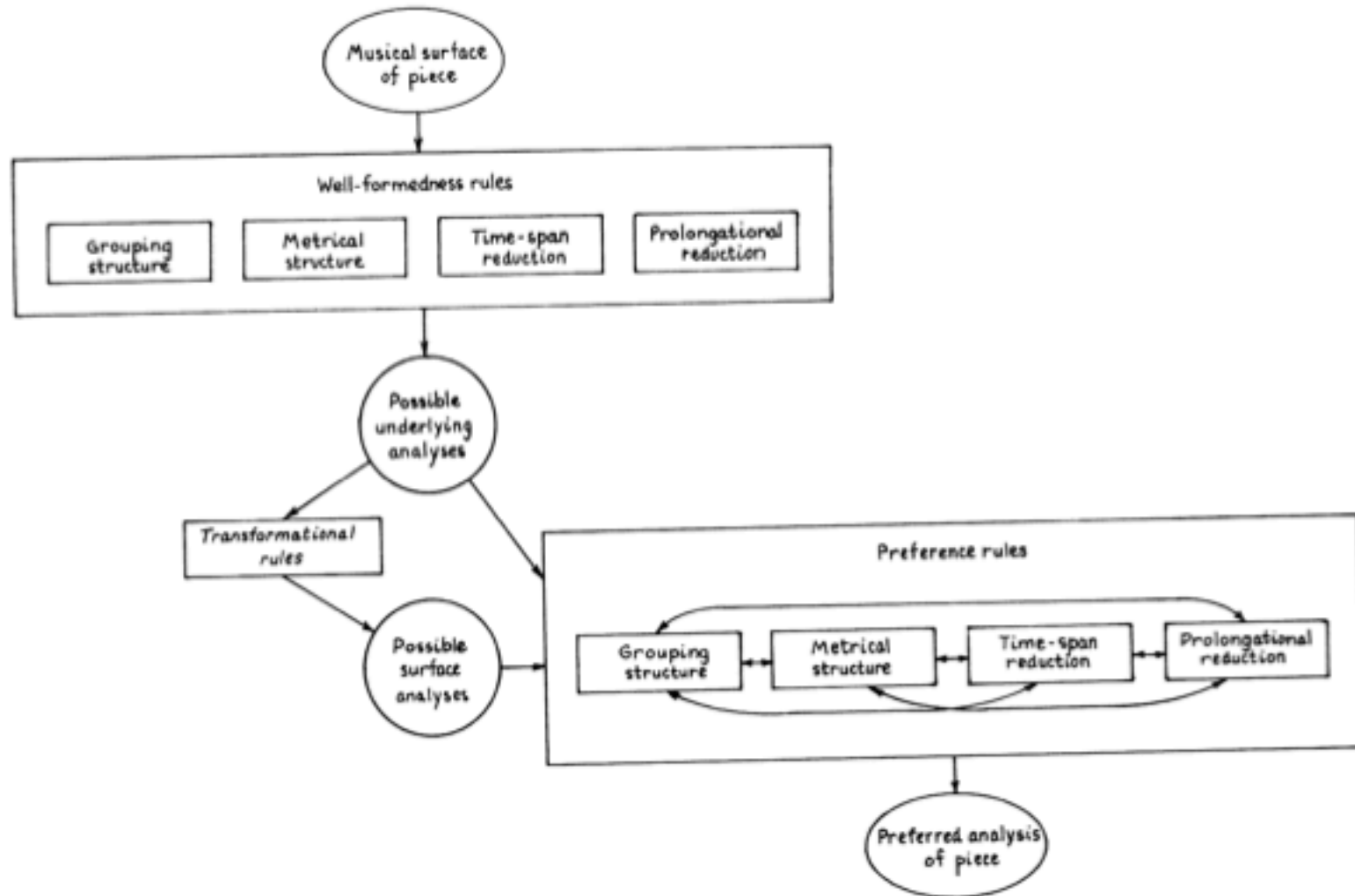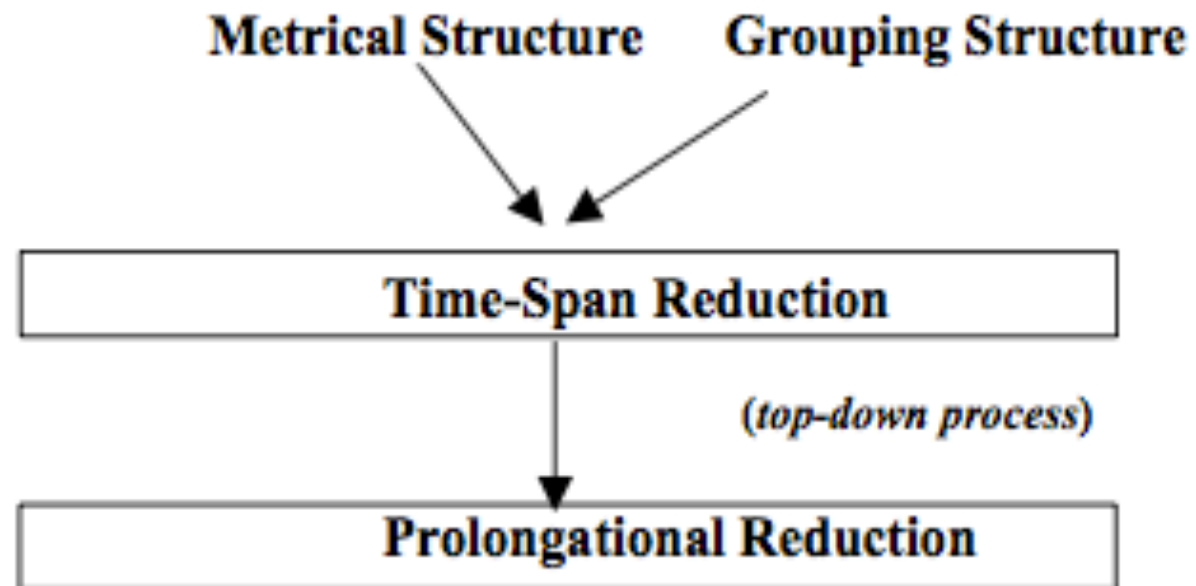
Figure 5.9 (A) A time-span reduction of the first two phrases of the children's song "Hush Little Baby." Shorter branches terminate on less important pitches, whereas longer branches terminate on more important pitches. (B) The lower staves show the dominant events at successively higher levels of tree structure. Modified from Large et al., 1995.
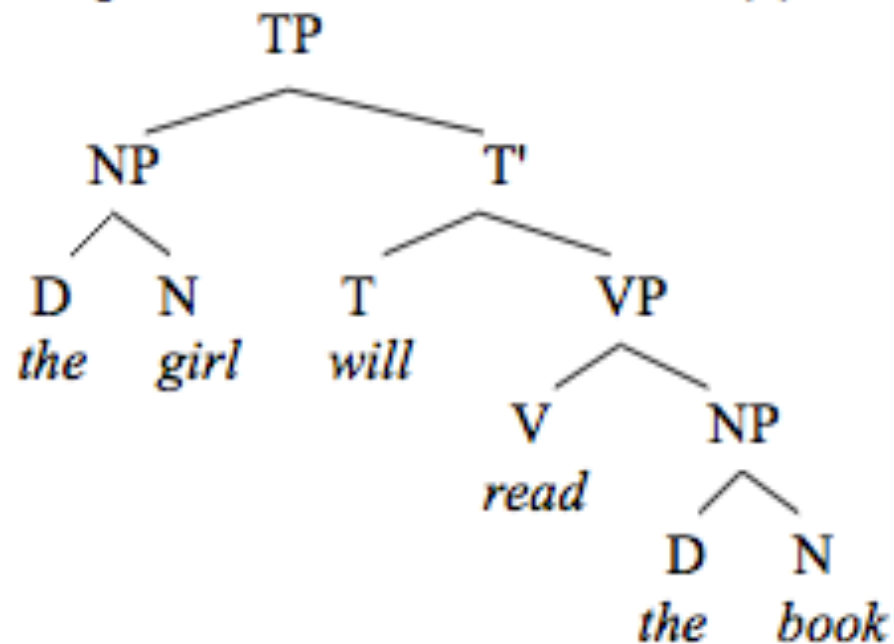
# GTTM's presentation of the model
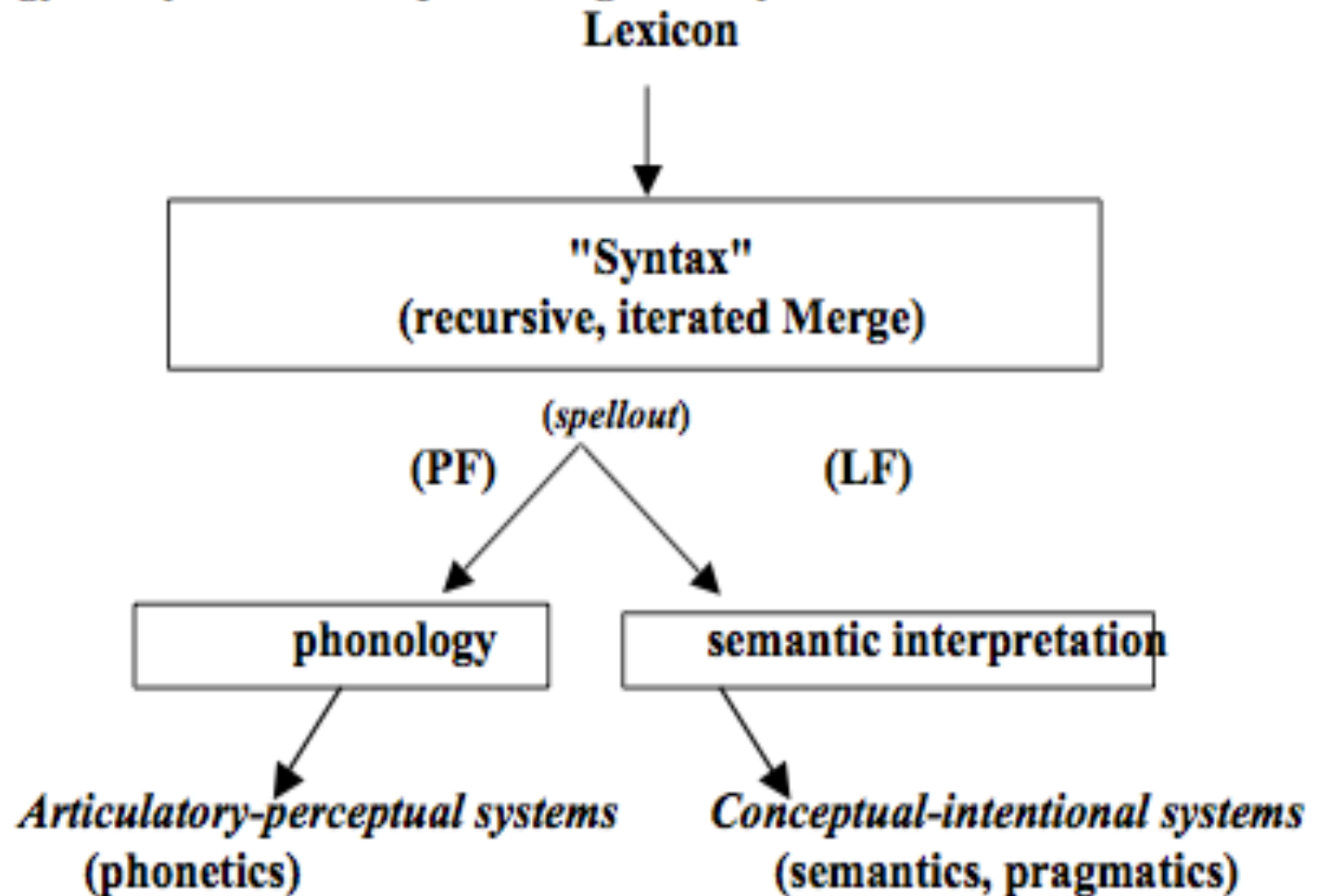
**(4)    A "practical boxology" for the GTTM model**

**Metrical Structure      Grouping Structure**

**Time-Span Reduction**

*(top-down process)*

**Prolongational Reduction**

**(5) Derivation by Merge of *The girl will read the book***

a. Form the set {the, book}                                      (lex. Item, lex. Item)
b. Form the set {read, {the, book}}                              (lex. Item, a)
c. Form the set {will, {read, {the book}}}                       (lex item, b)
d. Form the set {the, girl}                                       (lex. Item, lex. Item)
e. Form the set { {the, girl}, {will, {read, {the, book}}}}}  (d, c)

(6)     **Tree representation of derivation (5)**

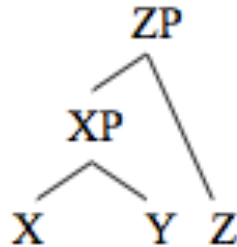(7)    **A boxology for syntax and adjacent cognitive systems**

Lexicon

$\downarrow$

| "Syntax"
(recursive, iterated Merge) |
|---|

(*spellout*)

(PF)      (LF)

| phonology |      | semantic interpretation |
|---|---|

*Articulatory-perceptual systems*
(phonetics)          *Conceptual-intentional systems*
(semantics, pragmatics)

# The syntax and prosody of language and music

- ## What aligns with what?
  - ### Musical TSR aligns with linguistic prosody.
    - The TSR representations of GTTM are similar in a number of respects to representations that linguists have proposed for prosodic structure, and the interpretation of both TSR and prosodic structures invokes notions of relative prominence.
  - ### Musical PR aligns with linguistic syntax.
    - The linguistic structures produced by Merge are similar to PR.
    - First, both can be understood as forming binary-branching, headed, acyclic directed graphs.
    - Second, the constituents of PR and linguistic syntax are headed, and thus encode structural relations between elements that are not necessarily string-adjacent.

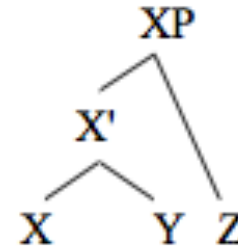# How linguistic trees and PR trees correspond
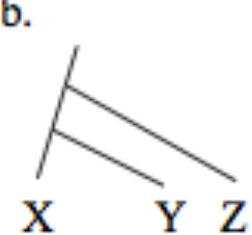
(10)  **X heads α; Z heads β**
a.  ZP    b.
  XP
  X   Y Z       X   Y Z

(11) **X heads α and β**
a.  XP    b.
  X'
  X   Y Z       X   Y Z

(12)  **Y heads α; Z heads β**
a.  ZP    b.
  YP
  X   Y Z       X  Y   Z

(13)  **Y heads α and β**
a.  YP    b.
  Y'
  X   Y Z       X  Y  Z
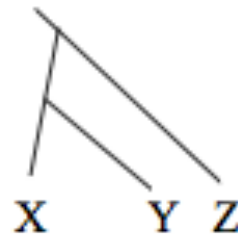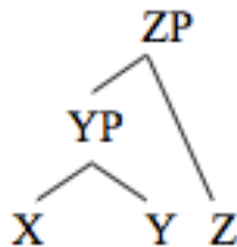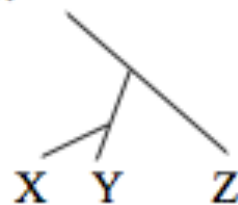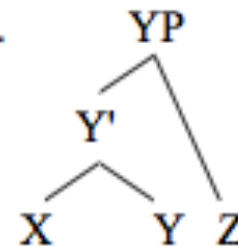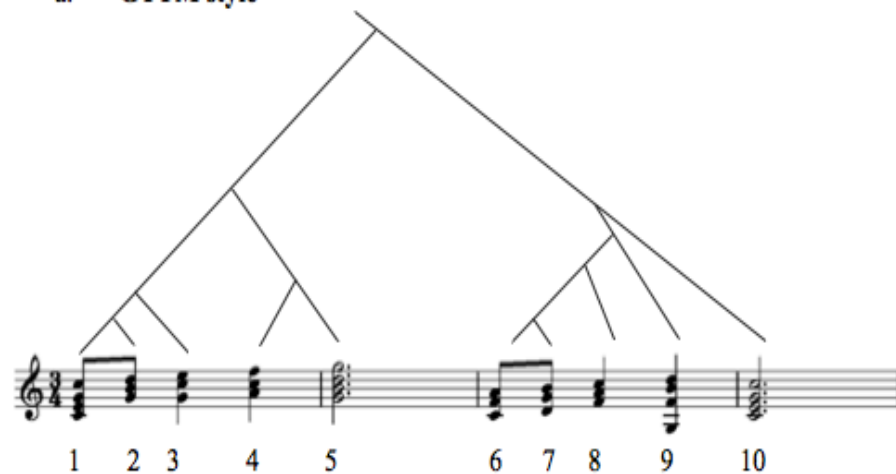
(14)  **PR structure for a toy melody**
  a.  **GTTM style**



  b.  **Linguist's style**
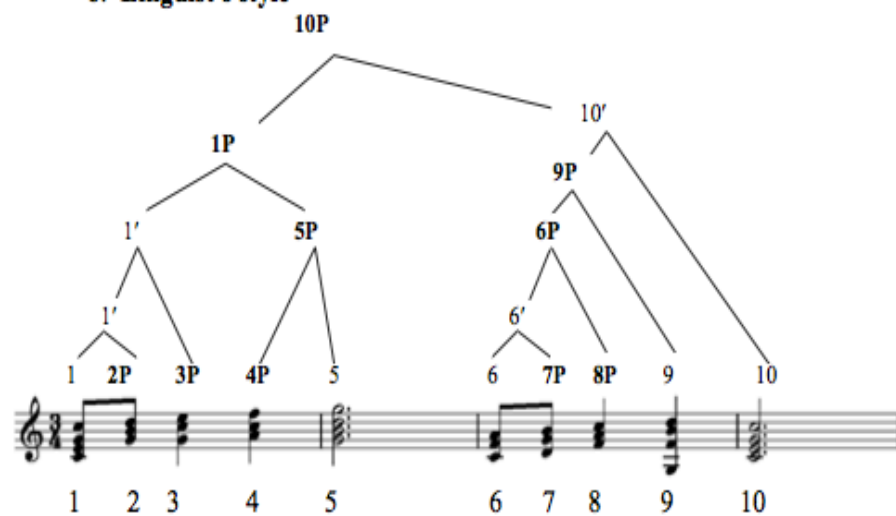
**(15)    GTTM-style tree corresponding to (6)**



|  D  |  N  |  T  |  V  |  D  |  N  |
|-----|-----|-----|-----|-----|-----|
| *the* | *girl* | *will* | *read* | *the* | *book* |
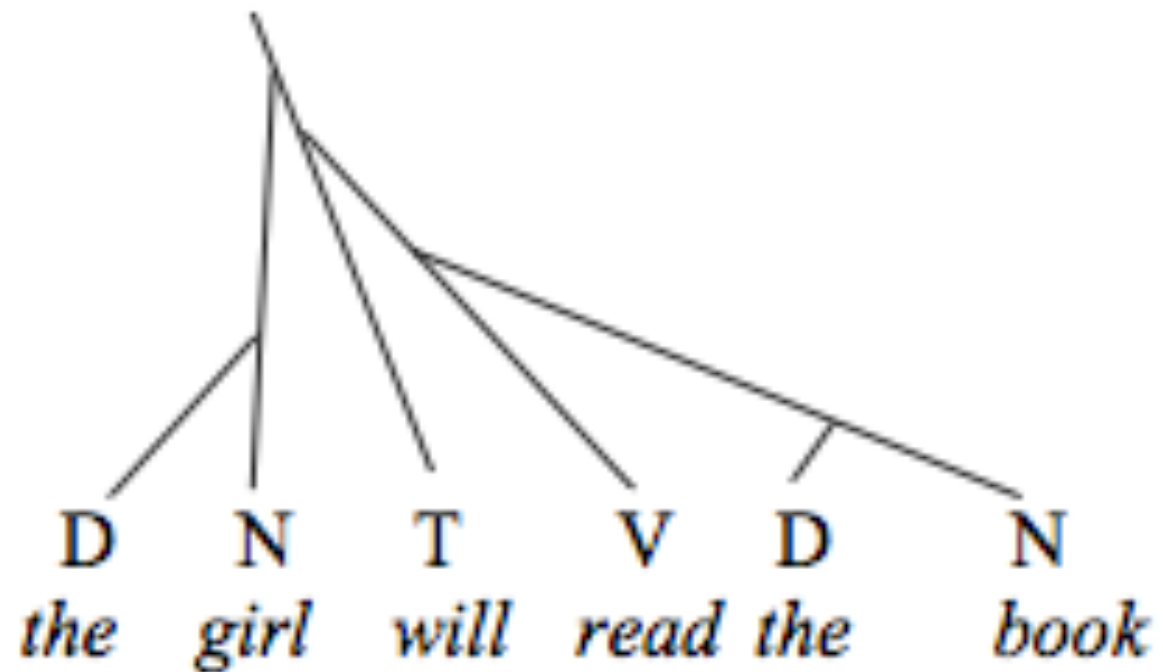
# Why both TSR and PR?

Note that in PR, the dominant chord y (G#-**E**-B) should be chosen over the subdominant chord x (**F#**-E-A) as head. ([Recording](#))



(18) **PR structure for Mozart piano sonata K. 331: GTTM notation**[16]
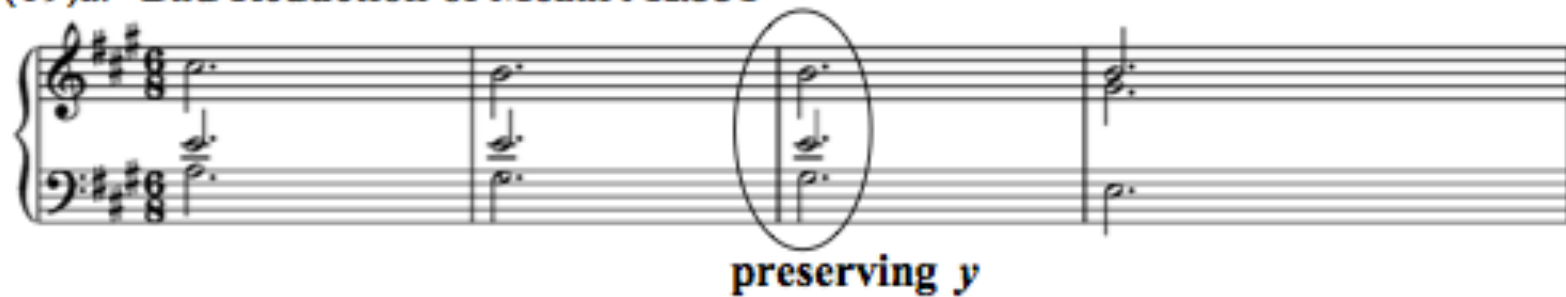(GTTM 231, ex. 9.17(b)[17]; Jackendoff (1987, 231))
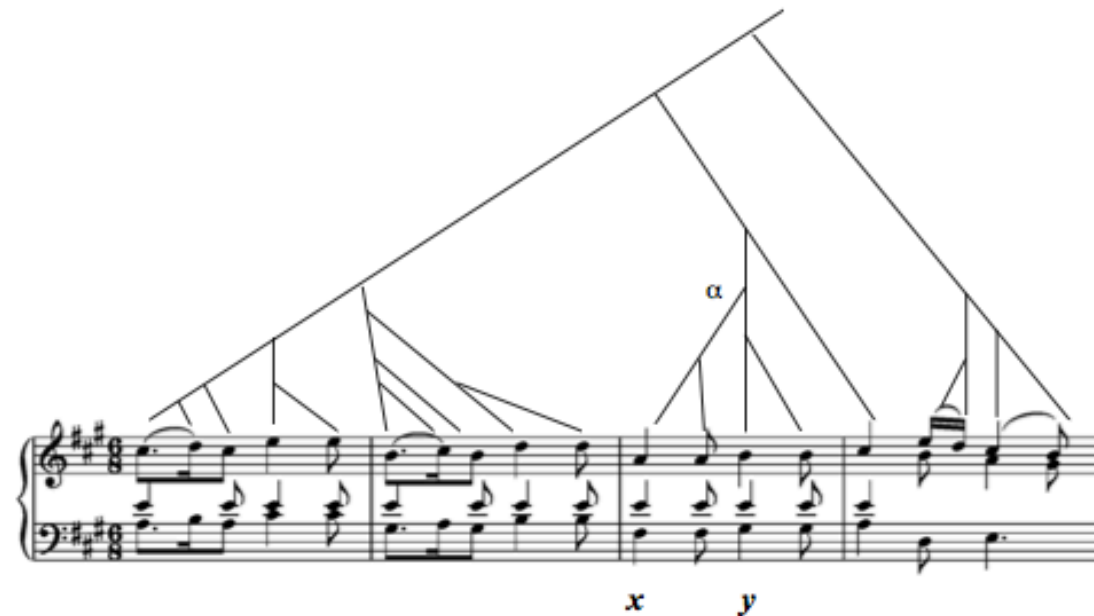
*low-level reduction:*

But this makes a bad reduction. ([Recording](#))



(19)a. **Bad Reduction of Mozart K.331**
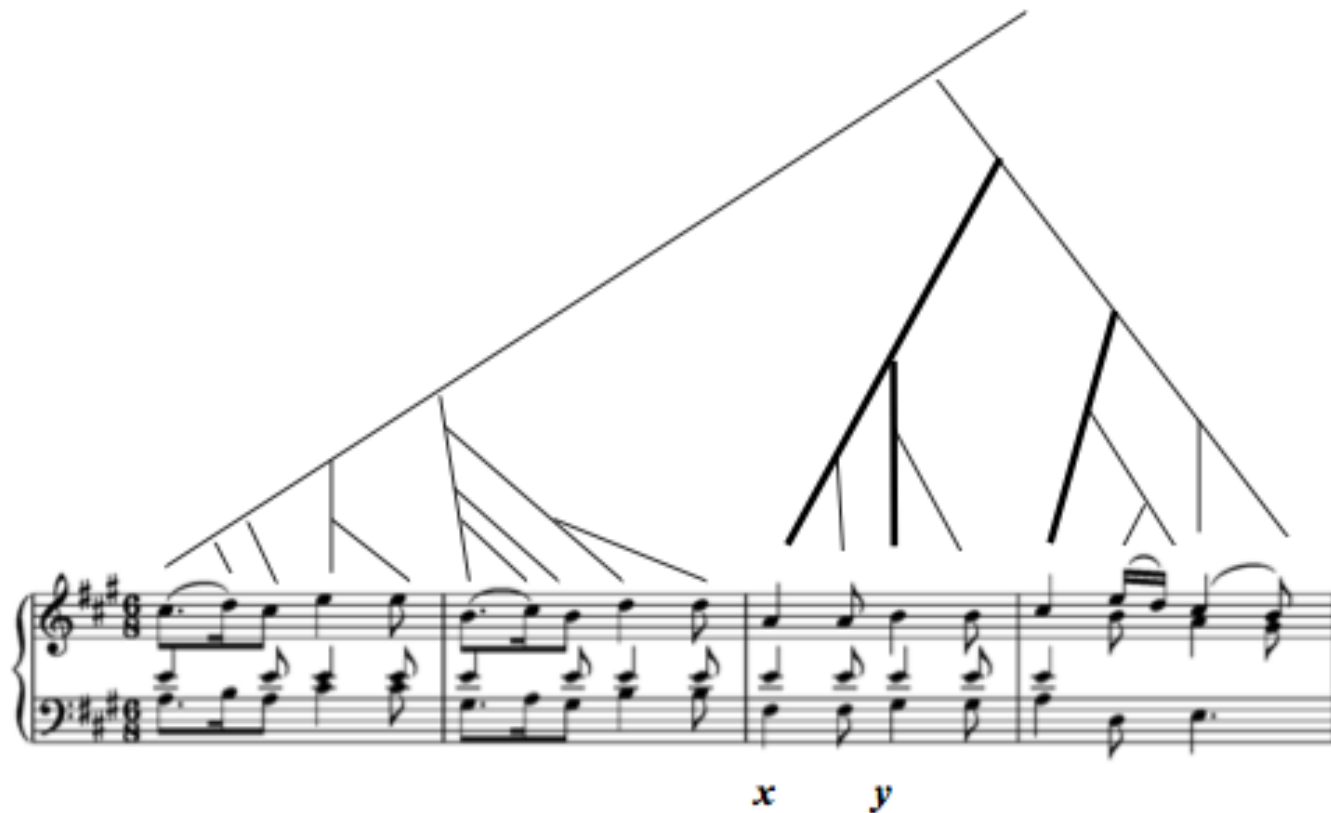
preserving *y*

b. **Good Reduction of Mozart K.331**

preserving *x*

PR defers to TSR in which the subdominant chord is stronger.



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

# Syntactic structure and prosodic structure in language

- Syntactic structure in language also defers to prosodic structure in determining the phonetic structure of a sentence.
  - This is [the cat that chased [the rat that ate [the cheese]]].
  - [This is the CAT] [that chased the RAT] [that ate the CHEESE].

# The Cadence as an analog in music of internal Merge in language

- "External Merge" puts two structures together that are up-to-that-point separate.

- "Internal Merge" puts a structure together with something that is already inside that structure.

  – It moves a structure from inside a constituent up and reattaches it.

    - Wh Movement
    - Head Movement

(37) *The boy wondered which book the girl will read.*

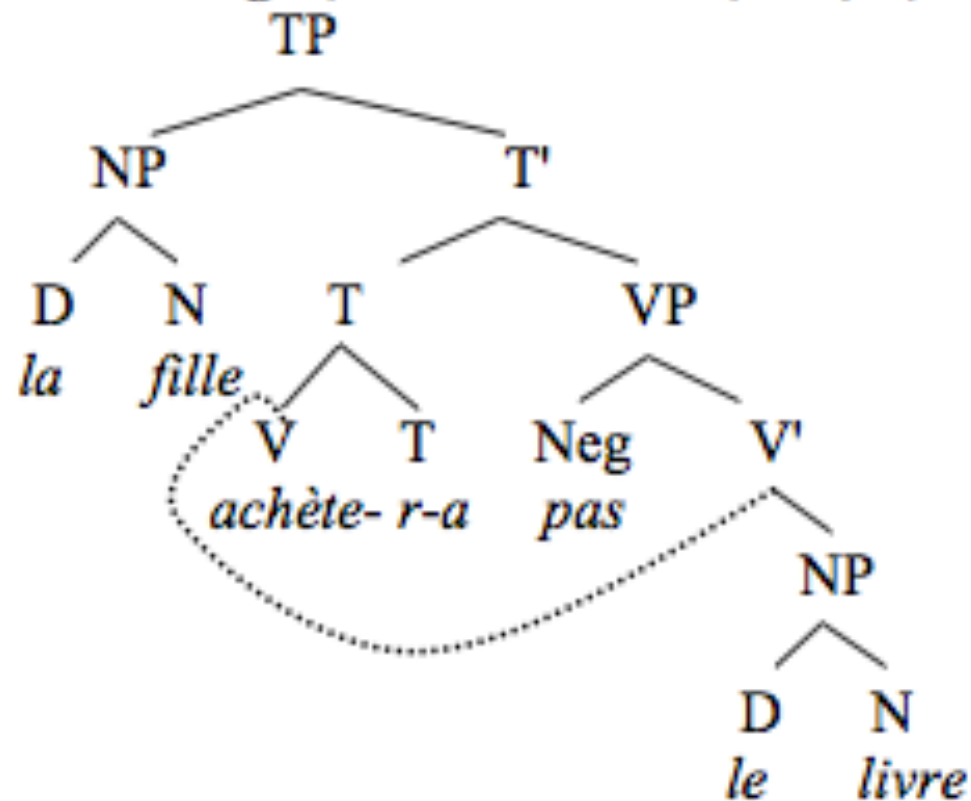(38)    **Internal Merge in the subordinate clause of (37)**
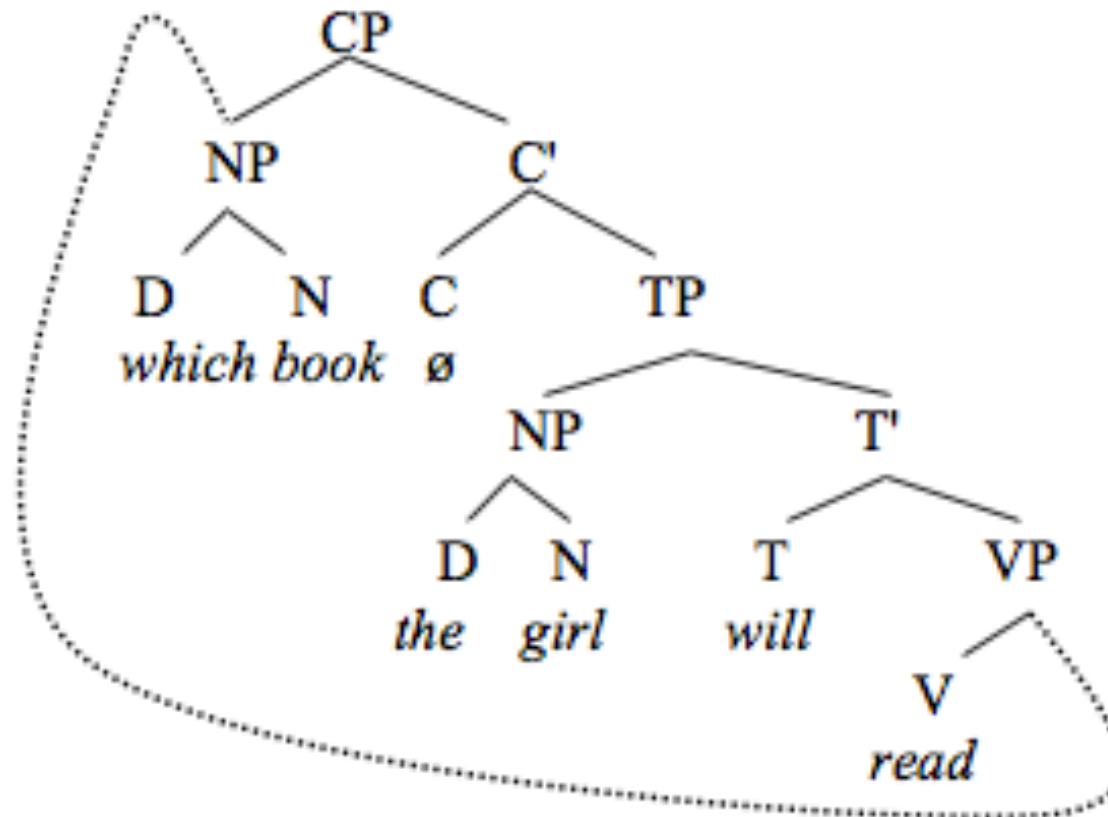
French:

*la   fille achetera   pas le  livre.*

the    girl   will.buy.3sg  not  the  book

'The girl will not buy the book'.

(40)    **Internal Merge (head movement) in (39)**

```
                        TP
              ┌──────────┴──────────┐
            NP                      T'
          ╱    ╲              ┌──────┴──────┐
        D       N            T             VP
       la     fille      ╱    ╲         ╱      ╲
                        V      T      Neg       V'
                    achète- r-a     pas       ╱   ╲
                                            NP
                                          ╱    ╲
                                        D       N
                                       le     livre
```

(49) **String-vacuous V-to-T head movement in Japanese**

| NP | N | Q | V | T |
|---|---|---|---|---|
| Mary-ga | ringo-o | 3-tu | kat- | -ta |
| Mary- NOM | apple-ACC | three-CLASS | buy | PAST |

'Mary bought John three apples.'
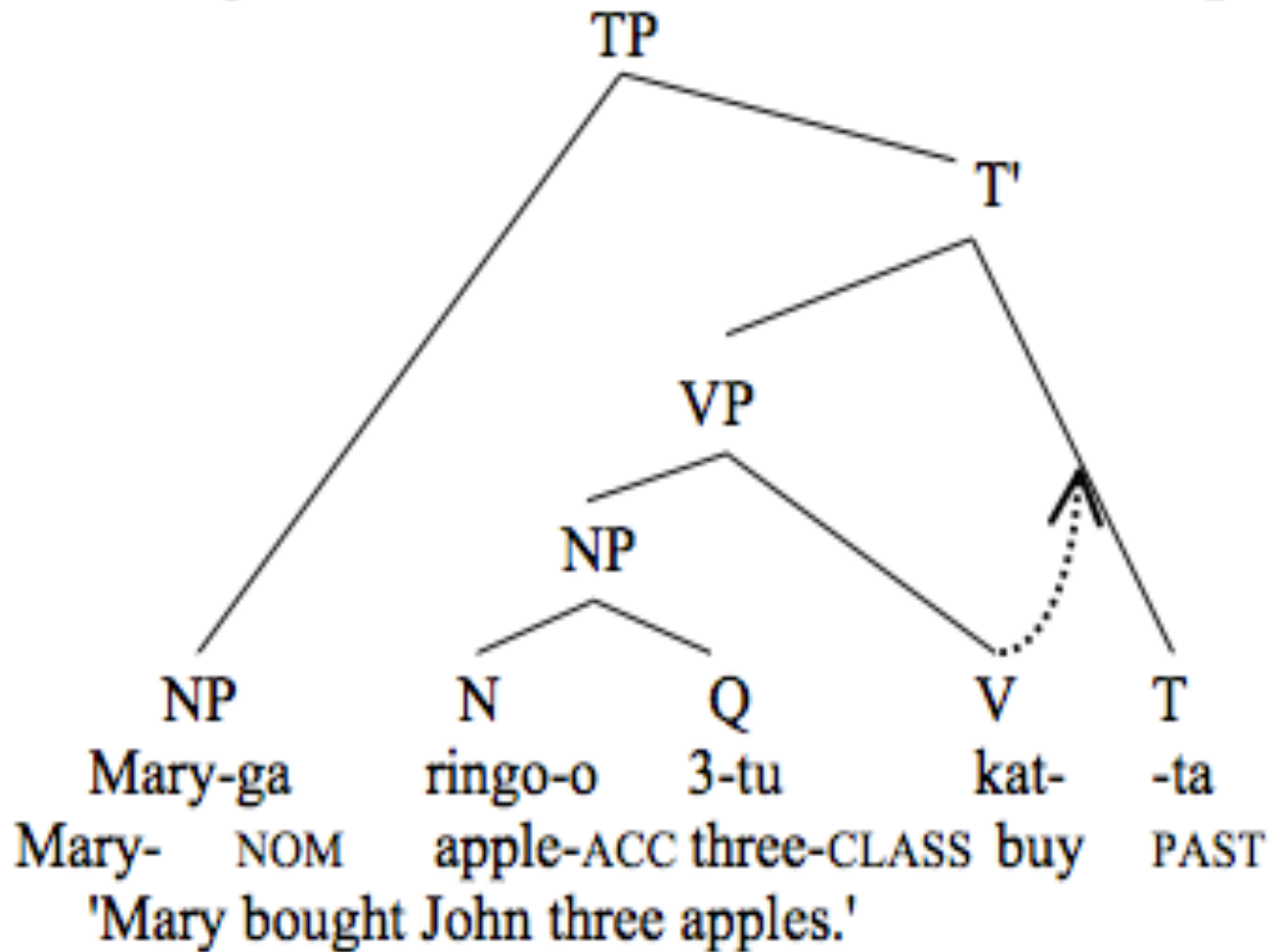
# String-vacuous head movement in music to create a cadence ([recording](recording))

(44) **Cadential Internal Merge of δ to τ (Head-Movement)**
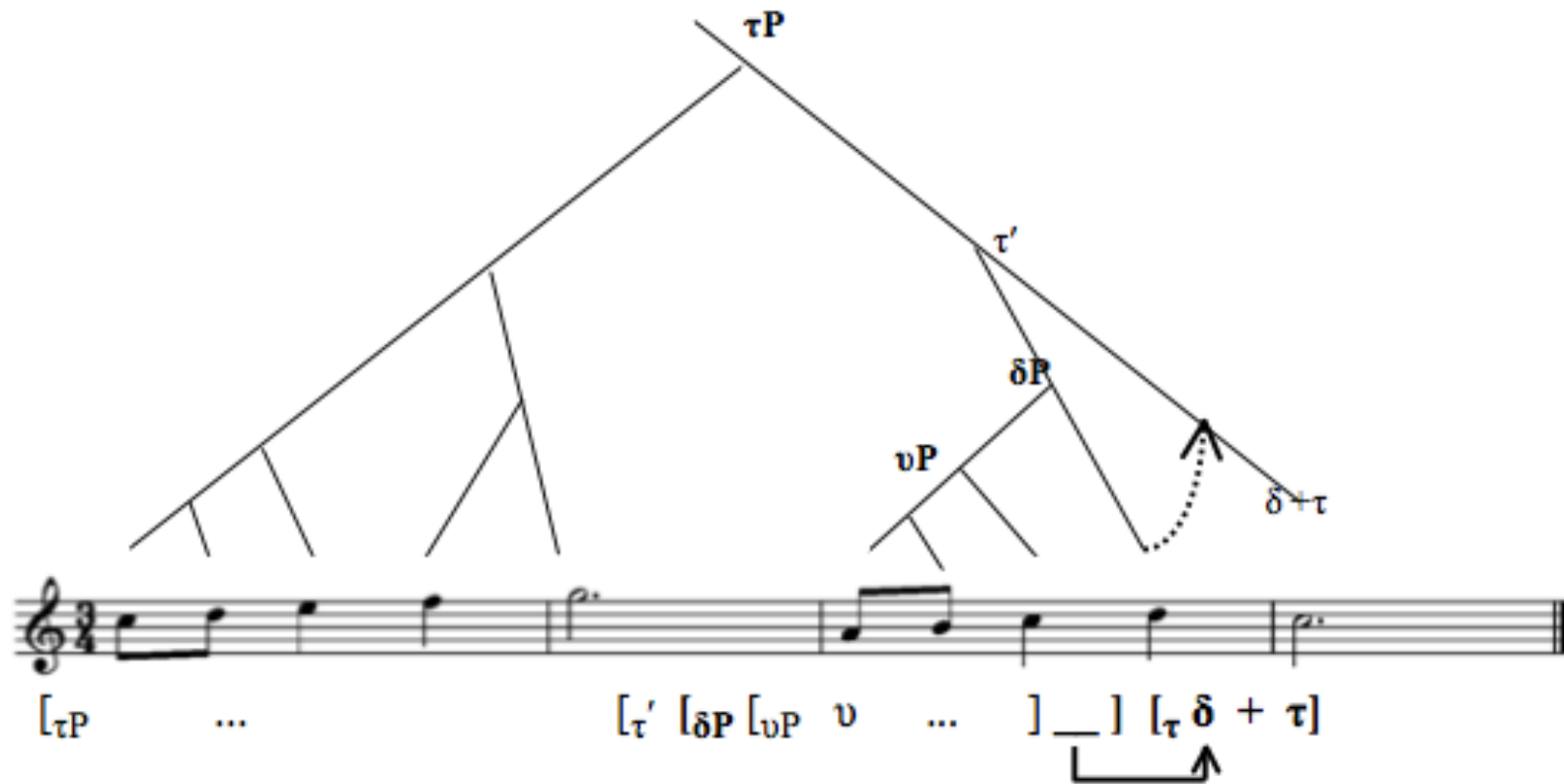
(43)    **Properties of Full Cadence**

A *full cadence* is a sequence of two chords[35] ($\delta$, $\tau$) drawn from the same pitch-collection (scale), where $\delta$ (the *dominant*) is a major triad whose root is higher by the interval of a fifth (= five scale steps) than $\tau$ (the *tonic*) — with the following additional properties:

  a.  $\delta$ must be **string-adjacent** to $\tau$, but at the same time...

  b.  ...$\delta$ may have a **normal set of syntactic dependents**, linearized normally. Thus $\delta$ also **heads its own phrase ($\delta$P).**

      The presence of at least one dependent of $\delta$ is exceedingly common, and is called a **subdominant,** which we will indicate with $\upsilon$ (GTTM, pp. 192 ff.).[36] It is typically built on the fourth or second step of the scale, but other possibilities are also found.

  c.  The participation of $\tau$ in a cadence ($\delta$, $\tau$) is crucial to **establishing the key as $\tau$.**

  d.  In PR, $\delta$ behaves as if it were structurally directly subordinate to $\tau$, even when its level of *prosodic* prominence (i.e. its RD number in TSR) should motivate a lower PR attachment site. GTTM calls this property **cadential retention.**

# The analog of semantics in music

- Cadences
  - As in linguistic head movement (V to T), cadences feature an adjacency requirement; they result in a tight coupling of two elements; the two elements still behave as if they are structurally independent; and the phenomenon is required in order to build a syntactic structure that can be interpreted.
- Establishment of a Key Domain
  - Head movement to establish a cadence is driven by the need for a piece of music to establish a tonic key.
  - A cadence is needed to satisfy a Principle of Full Interpretation, which thus applies in music as well as in language.
  - The syntactic structure PR thus feeds a component devoted to key interpretation. This Tonal Harmonic Component (THC) is the musical analog of semantics in language.