Quaderni di Linguistica e Studi Orientali- QULSO
Working papers in Linguistics and Oriental Studies

Firenze University Press
https://oajournals.fupress.net/index.php/bsfm-qulso

OPEN ACCESS

**Citation**: D. Pescarini (2020) Locality in music and syntax: a minimalist analysis of modulation. *Qulso* 6: pp. 161-171. doi: http://dx.doi.org/10.13128/QULSO-2421-7220-9698



**Data Availability Statement:** All relevant data are within the paper and its Supporting Information files.

**Competing Interests:** The Author(s) declare(s) no conflict of interest.

# Locality in music and syntax: a minimalist analysis of modulation

*Diego Pescarini*
CNRS, Université Côte d'Azur, BCL
(<diego.pescarini@univ-cotedazur.fr>)

*Abstract:*

The article provides an analysis of tonicization and modulation in western tonal music within the framework of minimalist syntax. In the spirit of Katz and Pesetsky's 2011 Identity Thesis for Language and Music, I argue that syntactic notions such as Agree and (Phase) Impenetrability may shed light on some aspects of music theory involving dependencies within local domains.

Keywords: *agree, locality, music, phases, syntax*

## *1. Introduction*

This article elaborates on some parallelisms between music theory and formal syntax, a line of inquiry inaugurated by Lerdahl and Jackendoff 1983. I will focus on *tonicisation* (Schenker 1906/1954: 256) and *modulation*, i.e. the harmonic processes whereby the *tonal centre* changes from one key to another. I will show that tonicisation and modulation hinge on syntactic-like *dependencies* that are constrained by *locality conditions*.

In syntax, locality conditions constrain syntactic dependencies such as agreement, *wh-* movement, binding, etc. For example, reflexive pronouns like *himself* must be bound by an antecedent belonging to the same clause, see (2)a *vs* (2) b. Roughly speaking, (2)b is impossible not because there are too many words between the reflexive and its antecedent, but because the dependency between the two crosses a structural barrier that impedes binding. Locality conditions are therefore sensitive to syntactic constituency (e.g. clause boundaries), while they are independent from processing issues such as short-term memory capacity.

(1) a. John$_i$ is impressed with *himself*$_i$
b. *John$_i$ asked Mary to help *himself*$_i$

In highlighting similarities between music and syntax, I intend to support Katz and Pesetsky's 2011 thesis in (2), although I will depart from their analysis of tonicisation and modulation.

(2) *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.

The article is organized as follows: §2 overviews some properties of tonal music; §3 reviews Katz and Pesetsky's analysis of *cadence*, *tonicisation*, and *modulation*; §4 suggests a revision of the analysis. §5 concludes.

## *2. Some features of tonal music*

Music perception relies on the categorisation of pitch events. In this respect, musical systems exhibit universal properties (Brown and Jordania 2011) such as:

(3) a. Use of discrete pitches rather than slides/portamentos.
b. Octave equivalence = unison choral singing in octaves.
c. Use of pitch sets = musical scales.

Hence, in all musical systems octaves are segmented into discrete intervals, yielding *scales* of pitch events (*notes*) ordered by fundamental frequency. The set of notes forming a scale is called *key* or *tonality* (henceforth: T). Although humans can distinguish up to 240 different pitches over an octave, musical traditions rely mainly on scales comprising five to seven tones (Gill and Purves 2008).

Notes can combine both 'horizontally', forming sequences called *melodies*, and 'vertically', forming simultaneous combinations called *chords*. Chords are sequenced according to *harmonic rules*, governing harmonic progressions within *musical phrases*. Rules are subject to variation across cultures and styles, but, like the rules of grammar, they are probably constrained by few invariable principles rooted in cognition.

Above all, the categorisation of pitch events, including chords, does not rely on absolute frequencies, but on scale degrees with respect to a reference pitch, called *tonic*. From now on, we will represent a tonality/key T as a set of scale degrees represented by Roman numerals; the *tonic* (t) is I in T:

(4) $T_{(t)}$ = {I, II, III, IV, V, VI, VII}

For instance, the note A[1], which correspond to a frequency of about 440Hz, is the degree VI in $T_{(C)}$ (i.e. in the tonality that has C as its tonic), V in $T_{(D)}$, III in $T_{(F)}$, I in $T_{(A)}$, etc.

[1] In the English-speaking world, notes are represented by letters (A, B, C, etc.), while other countries adopt *solmization*, i.e. a mnemonic attributing a syllable to each note C = *Do*, D = *Re*, E = *Mi*, F = *Fa*, G = *Sol*, A = *La*, B = *Si*.

In modern western tonal music, scales are normally formed by seven notes at intervals of a whole or half step.[2]

The most common patterns (termed *modes*) are the *major* and *minor* modes: in the former, half steps occur between the III and IV degree and the VII and I degree of the scale, while in the minor mode half tones occur between the II and III degree and between the V and VI grade (of the descending scale):

(5) Major mode: $I^n$ II III IV V VI VII $I_2$

Minor mode: $I^n$ II III IV V VI VII $I_2$

The distribution of whole/half steps is fixed, regardless of the frequency of t. Therefore, if the pitch of t changes, the pitch of the other notes forming the tonality must be adjusted. For instance, (6) shows the sets of notes forming the tonalities of C, D and F Major: the diacritics ♯ and b signal that the preceding note is raised or lowered of half step in order to obtain scales in the same mode:

(6) $T_{(C)}$ = {C, D, E, F, G, A, B, $C_2$}

$T_{(D)}$ = {D, E, F♯, G, A, B, C♯, $D_2$}

$T_{(F)}$ = {F, G, A, Bb, C, D, E, $F_2$}

Since the intervals within the scale are fixed, the three melodic contours (7) will be perceived as instantiations of the *same* melody in the tonalities $T_{(C)}$, $T_{(D)}$, and $T_{(F)}$, respectively. In fact, the four notes in (7)a-c have the same degrees (I III V I), but in different keys:

(7) a. C E G C
b. D F♯ A D
c. F A C F

Analogously, the perception of chord progressions depends on the degree of the chord's *root* (the root is the fundamental note of the chord; I will abstract away from the rules of harmony pertaining chord formation). For instance, a C major chord, which is formed by the notes {C, E, G}, is interpreted as V in $T_{(F)}$ because the root C is V in the key of F.

In conclusion, music categorization consists of a process of *interpretation* (Katz and Pesetsky 2011), whereby pitch events (tones, chords) are interpreted in relation to a reference pitch, called tonic (*t*). In our notation, interpretation maps notes (e.g. C, D, etc.) into degrees of a scale/key/tonality (e.g. I, II, etc.) in which *t* = I.

In the next sections, we will see that western tonal music allows *key change*: musical pieces can be therefore decomposed into *key domains* headed by a local tonic.

[2] Other types of scales can be obtained by dividing the octave in equal intervals of either a whole or half step. In the former case, the octave is divided into five intervals, yielding a hexatonic scale; in the latter we obtain a dodecaphonic scale.

### *3. Cadence, Tonicisation, and modulation*

A musical phrase is a string of pitch events that is perceived as an autonomous unit. A phrase is formed by subconstituents (*cells*) and it may be part of a larger *period*. Several factors contribute to defining musical phrases, including rhythm, melody, and harmony. With respect to harmony, musical phrases usually end with a *cadence*: a progression in which a chord conveying tension (e.g. V) resolves into a chord giving a sense of stability (e.g. I).

(8) a. $[_{\text{Phrase}}$ … V I] (cadence)
b. $[_{\text{Phrase}}$ … V I … ]

Notice that, to be perceived as a cadence, a V-I progression must occur at the *edge* of the phrase: schematically, V-I is a cadence in (8)a, but not in (8)b. In the following subsections I will introduce Katz and Pesetsky's 2011 analysis of the cadence. In §§3.2-3 I will elaborate on the two related components that yield cadence: tonicisation and key-domains.

### *3.1 Katz and Pesetsky's 2011 analysis*

Katz and Pesetsky 2011 argue for a parallelism between *head movement* in natural language and the cadence in western tonal music. In current syntactic theory, head movement is a kind of displacement resulting when a lexical head is moved without displacing the phrase it belongs to. For instance, the Italian verb *mangio* 'I eat' in (9) is the head of the Verb Phrase containing the object *piselli* 'peas'; in simple tenses, the verbal head is moved before the adjoined adverb *sempre* 'always', yielding the order:

(9) *mangio* sempre $[_{\text{VP}}$ ~~*mangio*~~ piselli]

Verb movement results in incorporation of the verbal head to a functional head (dubbed T for Tense), which encodes inflectional features. In compound tenses, T is spelled out by an auxiliary verb (see (10)a), whereas in simple tenses V moves (above the adverb) to incorporate T's features, see (10)b:

(10)a. [T [Adv $[_{\text{VP}}$ V NP]]]
*ho* *sempre* *mangiato* *piselli* 'I have always eaten peas.'

b. [V+T [Adv $[_{\text{VP}}$ V NP]]]
*mangio* *sempre* *piselli* 'I always eat peas.'

Katz and Pesetsky argue for a parallelism between head movement and cadence. They claim that, in a cadence between a chord δ and a tonic τ, the relationship established between δ and τ is comparable to the one holding between the two heads V and T in (10)b. The parallelism is supported by the following similarities, some of which will be discussed in the following subsections:

a. after head movement, the remnant phrase (e.g. the VP in (10)b) remains an independent phrase, which can be displaced autonomously regardless of the position of its head V; *mu-*

*tatis mutandis,* the basic key domain – our metaphorical "VP" – and the cadence – which is our metaphorical T+V complex head – are not necessarily adjacent, but they can be separated by various temporary tonic centres (more on this in §3.2);

b. head movement is obligatory; similarly, the participation of I in a cadence (e.g. V, I) is necessary to establishing the key (more on this in §3.3);

c. once the head has undergone head movement, it is pronounced string-adjacent to the higher head; the two end up tightly coupled, like the V and I chords in the full cadence.

In the remainder of the section, I will elaborate further on points (a) and (b). I will deal with the notion of Tonicisation first, which is instrumental in the analysis of modulation.

### *3.2 Tonicisation*

Given a tonality $T_{(t)}$, it is possible to introduce a chord X which does not belong to $T_{(t)}$. Extra-T chords (usually dubbed *altered* or *borrowed* chords) make the harmonic sequence richer, but their presence needs to be licensed by a process called *tonicisation* (Schenker 1906/1954: 256).

Tonicisation is a process whereby a chord that does not belong to the basic tonality is licensed by a local/temporary tonic. Take for instance a sequence of three major chords such as C, D, G in the tonality of C major. D major does not belong to $T_{(C)}$. However, the sequence is not ill-formed as the chord D precedes G, which acts as a temporary tonic for D as the chord D belongs to $T_{(G)}$:

(11) $[_{C}$ C D G … ]

By contrast, a chord such as Eb in the same position results in an illicit combination as Eb does not belong to the original tonality $T_{(C)}$ and it cannot be licensed by a temporary tonic as the chord of Eb does not belong to the secondary key $T_{(G)}$.

(12) $[_{C}$ C *Eb G … ]

In §3.1 I reported that Katz and Pesetsky hint at the possibility that the cadential formula may be preceded by a progression of *altered/borrowed* chords that do not belong to the basic key. Tonicisation does not establish a new key-domain (more on this below), but it consists of a temporary detour from the base tonality. Katz and Pesetsky notice that altered or borrowed chords often occur before the cadential formula, as shown in (13), in the same way in which adverbs or other adjuncts may occur between the VP and the complex head formed by V and T (cf. (11)b):

(13) [[[$_{C}$ … ] X Y Z ] V I ]

Basic tonality — altered/borrowed chords — cadence

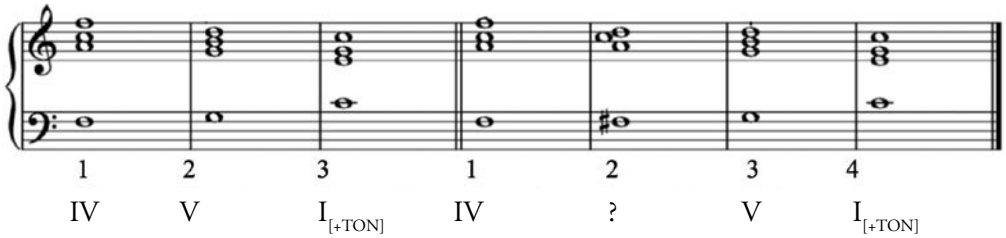
The fact that the cadence is separated from the body of the phrase in the base tonality is, according to Katz and Pesetsky 2011, the musical homologous of syntactic (head) movement.

### *3.3 Modulation*

Another common trait between the cadence and head movement is that both are obligatory. In particular, the cadence is necessary to establish a key domain and/or to establish a new tonic via *modulation.* Modulation differs from tonicisation (in §3.1): both yields tonic marking, but only the latter establishes a new key domain. To illustrate this point, let us compare a case of tonicisation with a full-fledged modulation.

Consider first the progression in (14)a and its extended variant in (14)b (corresponding to Katz and Pesetsky (58)): (14)a is formed by three chords belonging to the tonality of C (IV V I), whereas the latter contains an extra chord (in bar 2), which does not belong to $T_{(C)}$ (in is the same progression as in (11)).

(14) a. b.




One may argue that the altered chord in bar b2 is interpreted as V of an embedded key domain $T_{(G)}$, which is centered on the G chord of bar 3:

(15) $[_{C}$ … IV $[_{G}$ V $I_{+TON}$ ] $I_{+TON}$ ]

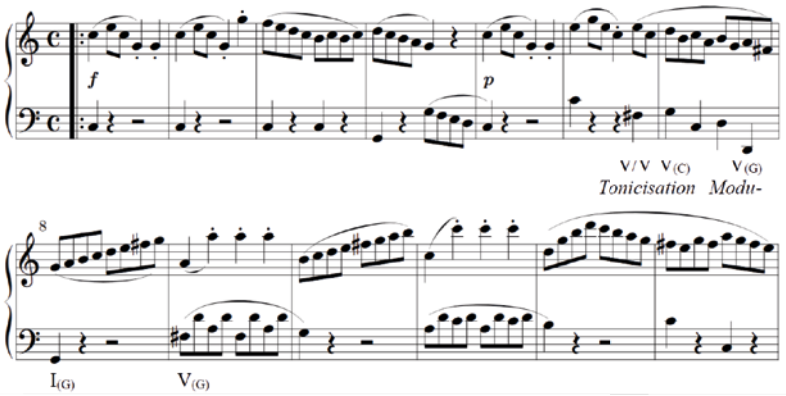
This conclusion, however, is not desirable as we perceive (14)b as a variant of (14)a. This amounts to saying that G is always interpreted as V in $T_{(C)}$ and no embedded key domain is established in (14)b. Instead, the altered chord in bar 2 is usually analysed as a *secondary chord* (noted as V/V), i.e. the V degree of the G chord, which in turn is the V degree of the basic tonality.

(16) $[_{C}$ … IV V/V V $I_{+TON}$ ]

No embedded key domain is established in (16): the G chord acts as a temporary tonic (cf. §3.1), but G does not establish its own Key Domain.

Conversely, when we modulate from a tonality (e.g. $T_{(C)}$) to another (e.g. $T_{(G)}$), we must establish a new key domain and all harmonic functions in the new domain must be assigned in the new key. The mechanism is illustrated in (17): at the beginning, the base tonality is $T_{(C)}$; then, at the end of bar 6 an altered D chord occurs and, as in (14b)/(16), D is interpreted as V/V in $T_{(C)}$ via tonicisation of G; in bar 7, however, the D chord is repeated, this time within a cadential formula (i.e. at the edge of the musical phrase). The combination of tonicisation and cadence yields modulation, i.e. a change of tonality from $T_{(C)}$ to $T_{(G)}$ that is permanent: starting from bar 8, all harmonic functions are assigned in the tonality $T_{(G)}$.

(17) Clementi, Sonatina op. 36 n. 1 (bars 1-13)




This simple example provides a clear comparison between tonicisation, which occurs between bars 6 and 7 (and does not result in key change), and modulation, which occurs between bars 7 and 8 (at the edge of the musical phrase), establishing a new key domain. Both tonicisation and modulation yield tonic marking, which, however, is not a necessary condition for establishing a key domain, cf. (15). In order to establish a key domain (via modulation), tonic marking must occur at the edge of a constituent ending with a cadence.

In the light of this conclusion, let us focus on Katz and Pesetsky's 2011 formal analysis of the cadence. In particular, I will focus on the relationship between three related concepts: cadence, key domain, and tonic-marking. In the following quote (from their §5.2.3), Katz and Pesetsky suggest that:

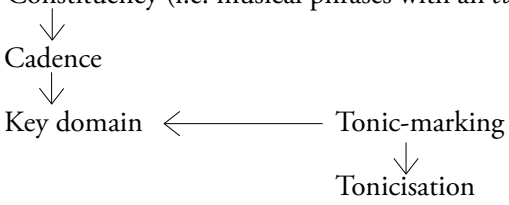
> cadential δ-to-τ movement has the function of **tonic-marking** τ, i.e. assigning it the feature [+TON]. When a head τ in a structure K is tonic-marked […], it has the consequence of allowing the terminal nodes of a particular subtree of K (determined by τ) to be understood as belonging to *the key of* τ.

According to the previous quote, the cadence assigns the feature [+TON] to a given pitch τ so that all others pitches belonging to the structure K are interpreted in the key $T_{(\tau)}$. However, this claim is too strong because also tonicisation assigns the feature [+TON] to a given pitch. In Katz and Pesetsky's 2011 words in (18), being marked [+TON] is a necessary, but not sufficient condition to establish a key domain:

(18) Key Domain
*Optional*: A node marked [+TON] is a key-domain.

In order to establish a key domain, a second condition must be met, namely that tonic-marking occurs in a cadence (recall that a cadence is a progression in which a chord conveying tension (e.g. V) resolves into a chord giving a sense of stability (e.g. I) at the edge of a phrase). The logical relationships between the concepts introduced so far is eventually schematized in (19): modulation results from two independent conditions, tonic-marking and cadence; the latter in turns depends on constituency. When tonic-marking takes place outside of a cadence, no key domain is established (§3.2):

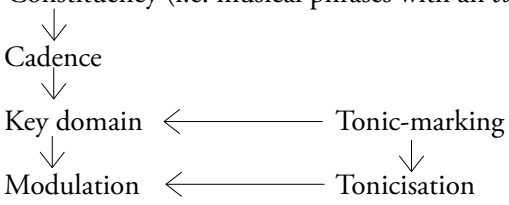
(19) Constituency (i.e. musical phrases with an *edge*)




Following (19), the cadence *per se* does not have "the function of tonic-marking τ", although it has the power of establishing a key domain by promoting a node marked [+TON] to the role of tonic of a key-domain.

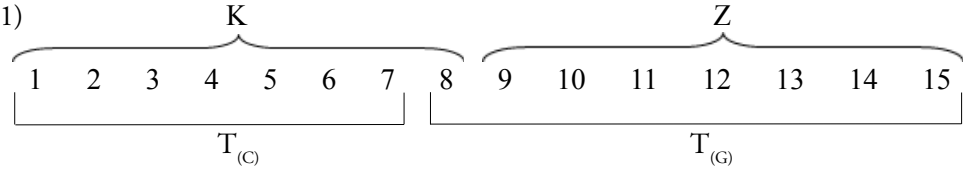
Given this state of affairs, we can define modulation as a process of tonicisation that takes place in the context of a cadence: if a new tonic is introduced at the edge of a phrase, the new tonic will be will become the centre of the newly established key domain:

(20) Constituency (i.e. musical phrases with an *edge*)




The latter point is illustrated (21), which is an analysis of the excerpt (17). In (17), bars 1-8 form a phrase K, whereas bars 9-15 form another phrase Z. The first seven bars of K are in the tonality $T_{(C)}$; then the cadence between bar 7 and 8 establishes a new key domain – $T_{(G)}$ – which is the tonality of the following phrase Z:

(21)




What is crucial is that the key domain $T_{(G)}$ crosses the phrase boundary between K and Z. This contradicts (my understanding of) Katz and Pesetsky's hypothesis, according to which the cadence determines the harmonic functions of the phrase *it belongs to*, namely K (see the above quote from Katz and Pesetsky). Rather, the cadence yields a modulation that creates a key domain that is formed by the right edge of K and the following phrase Z. In my opinion, this misalignment between phrases and key domains requires a partial reformulation of Katz and Pesetsky's analysis, which will be discussed in the next section.

*4. Towards a revision of Katz and Pesetsky's account*

This section aims to revise the analysis of tonicisation (§3.2) and modulation (§3.3) in order to meet two desiderata:

- divorce the definitions of tonicisation (tonic marking) and cadence;

- revisit the notion of cadence in order to provide a better analysis of the relationship between phrases and key domains.

§4.1 argues that tonicisation consists of an Agree relation between an uninterpretable chord (the Probe) and an interpretable chord (the Goal); §4.2 suggests that modulation results from an Impenetrability Condition, which prevents successive pitch events from being interpreted according to a previous tonic centre.

### *4.1 Tonicisation as Agree*

As mentioned in §2, *interpretation* is a process assigning *harmonic functions* (expressed by Roman numerals, e.g. I, II, V, etc.) to pitch events with respect to a key. If a chord does not belong to the base key, it is *uninterpretable* as it cannot be assigned a harmonic function. For instance, given the progression C D G in $T_{(C)}$, D is uninterpretable (*u*) as it has no harmonic function in $T_{(C)}$:

(22) Chords: C D G
Functions: $iI_{(C)}$ $u_{-(\ )}$ $iV_{(C)}$

In current minimalist theory, an uninterpretable element acts as a Probe searching for an interpretable Goal in a local domain. In the case of music, the goal is a nearby – though not necessarily adjacent – interpretable chord which may act as a temporary tonic for the probe (see §3.1). In (22), a licit probe-goal relation can be established between *u*D and G as the latter acts as a temporary tonic because D belongs to the (secondary) key $T_{(G)}$. Given the presence of G, the chord D is eventually interpreted as a second-grade harmonic function: the notation in (23) shows that the uninterpretable chord D becomes interpretable when it is *valued* by the harmonic function of the chord it agrees with:[3]

(23) Chords: C D G
Functions: $iI_{(C)}$ $iV_{(V)}$ $iV_{(C)}$
Agree

In conclusion, Agree allows a temporary deviation from a given tonality without establishing a new key domain. It seems to me that, by deriving tonic marking from a mechanism of agree,

[3] An anonymous reviewer pointed out that Chomsky's Agree consists of matching under feature identity (Chomsky 2000: 122). However, matching refers to a system in which various kinds of features co-exists and probe-goal relations can be established only between objects with matching features. Conversely, music relies only on harmonic functions: I, III, IV, etc. In a system like this, matching is always ensured since it is based on a sole type of feature.

This leads me to address another remark made by the same reviewer, who noticed that syntax is standardly ordered by dominance, while music (like phonology) is necessarily ordered by precedence. This, however, is not entirely true. Dominance plays a fundamental role in music computation, as argued extensively by Lerdahl and Jackendoff 1983. In the same line, Katz and Pesetsky's 2011 analysis of modulation entails movement within a hierarchical structure. However, it is true that music displays fewer types of long-distance dependencies than language. In my opinion, this is partly due to the fact that interpretation in music hinges on a single set of harmonic functions (I, II, III), whereas in language it results from various kinds of features that must be matched. If no intervening feature occurs, matching of syntactic feature may result in long-distance dependencies. In music, conversely, matching always takes place between adjacent/close pitch events because, having a sole set of harmonic functions, long-distance dependencies of the linguistic kind are disfavoured.

we eventually divorce tonicisation from the cadence. The former results from a formal mechanism that, in music as well as in grammar, turns uninterpretable elements into interpretable ones. For this reason, it seems to me that Agree provides a suitable homologous of tonicisation in syntactic theory.

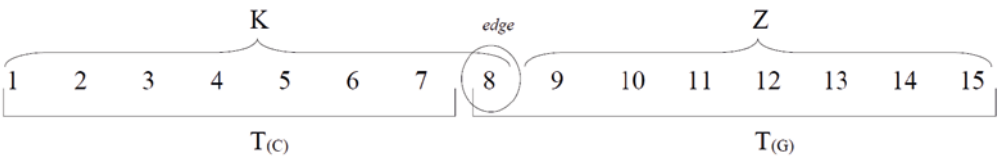
*4.2 Modulation as Phase Impenetrability*

Syntactic dependencies, which are ultimately reduced to probe-goal relations, are constrained within local domains. A principled account of local domains is provided by Chomsky's *Phase theory* (1999, 2001). Phase theory assumes that syntactic computation proceeds in *phases*: once a syntactic subtree is built, it is spelled out, i.e. it is sent to the semantic and phonetic interfaces in order to be mapped into phonological and conceptual structures, respectively. When a phase is sent to the interfaces, its inner structure is no longer available for further computation, while its outer layer is still visible from the successive phase. This hypothesis is captured by the *Phase Impenetrability Condition* (PIC):

(24) In a structure $[_{ZP}$ Z…$[_{HP}$ α [H YP]]], where Z and H are phase heads, the domain of H is not accessible to operations at ZP; only H and its edge are accessible to such operations

Let us suppose that musical structures, like syntactic structures, are organized into phases, which define impenetrable domains. If this analysis is on the right track, one can eventually suggest that notions such as key-domain and cadence are better analysed in terms of phasehood, which is independent from the algorithm building phrase structure/constituency. To illustrate the hypothesis, let us resume the analysis of Clementi's excerpt in (17): recall that the first 8 bars of Clementi's *sonatina* form a phrase K ending with a cadence; the *edge* of K (containing the cadence) forms a key domain in $T_{(G)}$ with the next phrase (bars 9-15).

(25)




This recalls Chomsky's discussion (1999: 10) concerning the spell-out domain in a structure of the following type:

(26) $[_{ZP}$ Z … $[_{HP}$ α [H YP]]]

Chomsky argues that "H and its edge α in (8) [= our (26)] belong to ZP for the purposes of Spell-out, under PIC". Thus, at least in its original formulation, Chomsky hints at the possibility that the edge of a phase forms a Spell-out domain with the superordinate phase ZP. It seems to me that the same intuition may provide a suitable account for the (mis)alignment between musical phrases and Key domains: a phrase K ends with a cadence, which may establish a new key domain (to which the next phrase Z belongs). The analysis of cadences as edges, which is compatible with Katz and Pesetsky's 2011 analysis in terms of head movement, provides a promising account of the mapping between constituent structures ("phrases") and key domains ("phases").

Most importantly, the PIC in (24) provides a sound explanation of modulation: in fact, the key domain preceding the cadence becomes impenetrable from the successive phase. Thus, any chord belonging to Z in (25) will be interpreted with respect to the new tonic established in bar 8, while no probe-goal relation can be established with any chord belonging to the previous key domain.

## *5. Conclusions*

The present paper has addressed the notions of *tonicization* and *modulation* in the framework of minimalist syntax. In the spirit of Katz and Pesetsky's 2011 *Identity Thesis for Language and Music,* I argued that syntactic notions such as Agree and Phase (Impenetrability) may shed light on some aspects of music theory involving dependencies within local domains.

I suggested a partial revision of Katz and Pesetsky's analysis by divorcing tonicisation from the cadence and by redefining the role of the cadence in establishing key domains. I argued that tonicisation consists of an Agree relation turning an uninterpretable chord into an interpretable pitch event by sharing the harmonic function of a nearby interpretable element. Lastly, I suggested that modulation results from a locality constraint reminiscent of Chomsky's 1999 *Phase Impenetrability Condition*, which prevents the interpretation of pitch events in a previous key if tonicisation takes place at the edge of a phase.

## *References*

Brown, Steven and Joseph Jordania. 2013. "Universals in the World's Musics." *Psychology of Music* 41, 2: 229-248. <https://doi.org/10.1177/0305735611425896> (06/2020).

Chomsky, Noam. 1999. "Derivation by Phase." *MIT Occasional Papers in Linguistics* 18. MITWPL, Department of Linguistics and Philosophy, MIT, Cambridge.

Chomsky, Noam. 2000. "Minimalist Inquiries." In *Step by Step: Essays on Minimalism in Honor of Howard Lasnik*, ed. by Roger Martin, David Michaels, and Juan Uriagereka, 89-155. Cambridge: MIT Press.

Gill, Kamraan Z., and Dale Purves. 2009. "A Biological Rationale for Musical Scales." *PLoS ONE* 4, 12: e8144. doi: 10.1371/journal.pone.0008144.

Katz, Jonah, and David Pesetsky. 2011. "The Identity Thesis for Language and Music." <lingbuzz/000959> (06/2020).

Lerdahl, Fried and Ray Jackendoff. 1983. "A Generative Grammar of Tonal Music." Cambridge: MIT Press.

Schenker, Heinrich. 1906/1954. *Harmony* (English edition by Oswald Jonas). Chicago: University of Chicago Press.