Chapter 3

Rethinking remerge: Merge, movement and music

Hedde Zeijlstra
Georg-August-Universität Göttingen

In an influential paper, Katz & Pesetsky (2011) present the identity thesis for language and music, stating that "[a]ll 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." Katz & Pesetsky argue that, just like syntactic structures, musical structures are generated by (binary) Merge, for which they provide a number of arguments: for instance, musical structures are endocentric (each instance of Merge in music, just like in language, has a labelling head). They also argue that movement phenomena (i.e., the application of Internal Merge) can be attested in both language and music. While fully endorsing the view that musical structures are the result of multiple applications of External (binary) Merge, this paper argues that the arguments in favour of the presence of Internal Merge in music are at best inconclusive and arguably incorrect. This is, however, not taken as an argument against the identity thesis for language and music; rather, I take it to follow from it: the identity thesis for language and music reduces all differences between language and music to its basic building blocks. If the application of Internal Merge in natural language is driven by uninterpretable features (cf. Chomsky 1995; 2001; Bošković 2007; Zeijlstra 2012) that are language-specific and not applicable to music (the reason being that only building blocks that are pairings of sound and meaning can be made up of interpretable and uninterpretable features), the direct consequence is that Internal Merge cannot be triggered in music either.

Hedde Zeijlstra. 2020. Rethinking remerge: Merge, movement and music. In András Bárány, Theresa Biberauer, Jamie Douglas & Sten Vikner (eds.), *Syntactic architecture and its consequences II: Between syntax and morphology*, 43–66. Berlin: Language Science Press. DOI: 10.5281/zenodo.4280631

1 Introduction: External and Internal Merge in language and music

Since Chomsky (1995), the operation Merge has been taken to be the primary structure-building operation in natural language. In current minimalism, syntactic movement is, moreover, considered a special instance of Merge (Internal Merge), which applies to a particular syntactic object and a part thereof (cf., inter alia, Chomsky 2005). In this sense, Internal Merge is different from External Merge, where the two input objects do not stand in an inclusion relation.

However, natural language is not the only cognitive domain where Merge is said to be a structure-building operation. As has been claimed in Lerdahl & Jackendoff (1983) and, more recently, in Katz & Pesetsky (2011), music is also a cognitive domain where structures can be taken to be generated by means of an operation like Merge. If musical structures are indeed generated by means of Merge and if movement is a special instance of Merge, the question arises whether music exhibits movement effects as well. After all, why could Internal Merge not apply in music if it can apply in natural language?

In order to account for the differences and similarities between language and music, Katz & Pesetsky (2011) entertain their so-called *identity thesis for language and music*, which states that:

> [a]ll 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. (Katz & Pesetsky 2011: 3)

For Katz & Pesetsky, this means that Merge should be equally effective in natural language and music and that therefore music is indeed expected to exhibit both External and Internal Merge effects. In their paper, they identify particular musical patterns that they take to reflect movement in music.

However, one may wonder whether it is correct to assume that identity thesis for language and music entails that both External and Internal Merge should apply in music. As I will argue in this paper, it all depends on what triggers Internal Merge in the first place. Internal Merge differs from External Merge in the sense that Internal Merge does not have to take elements from the numeration into the syntactic structure. If every element in the numeration needs to end up in the syntactic structure, it follows immediately that every element present in the numeration needs to undergo External Merge. But why would particular elements be required to undergo Internal Merge as well?

Following a longstanding tradition in syntactic theory, I assume that Internal Merge is triggered by so-called uninterpretable formal features – formal features that need to stand in a particular configuration with their interpretable counterparts. If that is the case, the question arises as to whether such movement-triggering features can also be attested in music. I argue they do not.

According to the identity thesis for language and music, all differences between music and language should reduce to differences in their building blocks: for Katz & Pesetsky, arbitrary pairings of sound and meaning in the case of language, and pitch classes and pitch-class combinations in the case of music. Let's focus in more detail on each type of building blocks.

Lexical items are generally thought to consist of three types of features: phonological features, syntactic or formal features, and semantic features. Phonological features are only interpretable or legible for the sensori-motor system; semantic features are only interpretable or legible for the conceptual-intentional systems; and syntactic or formal features are interpretable or legible for neither of them. In that sense, linguistic building blocks can be said to be multi-modular, not mono-modular.

Things are different when it comes to musical building blocks. One dimension in which the architecture of music is much different from that of natural language is that musical structures are not subject to compositional semantic interpretation in the sense that the meaning of a musical structure – to the extent it has any (see, for instance, Schlenker 2016 and references therein for discussion) – follows compositionally from the meaning of the parts it consists of and the way these parts are structured. While linguistic objects are built of elements that form sound-meaning pairs, the musical objects are not. Musical building blocks are mono-modular building blocks. Mono-modular building blocks are building blocks that are all interpretable or legible for the same module, in this case the sound side of music. And even if it turns out that pitch classes and pitch-class combinations are not the only available building blocks in music (and other building blocks are available as well, either inside or outside Western tonal music), those building blocks will still belong to the same sound module.

Mono- vs. multi-modularity is then a main characteristic of the differences between musical and linguistic building blocks. Now, under the view that the application of Internal Merge is indeed driven by the need of so-called uninterpretable features to be checked by their interpretable counterparts, it follows immediately that Internal Merge can only be triggered by features present on linguistic building blocks, not on musical building blocks. The reason is that uninterpretable features are defined as elements that are not part of the set of semantic features, but require a particular checking (or valuation) relation with a feature that does

belong to this set. As a consequence, no uninterpretable feature can be acquired without the presence of a semantic counterpart (see Brody 1997; Svenonius 2007; Zeijlstra 2008; 2012). But if that is correct, uninterpretable features, by definition, can only be part of building blocks that are not mono-modular. In fact, in any cognitive system whose output is not defined in terms of pairs of elements belonging to different cognitive modules (in the way that linguistic output is defined in terms of sound-meaning pairs), features that denote dependencies on elements belonging to different modules cannot exist.

If that is the case, the identity thesis for language and music should actually predict that, to the extent that Internal Merge can only be triggered by uninterpretable formal features, it can never apply to pieces of musical structure and that therefore instances of movement are expected to be absent in music.

In this article, I first further elaborate the claim that (properties of) uninterpretable features are the trigger for syntactic movement (§2). Then, in §3, I discuss Katz & Pesetsky's claim that music does not only exhibit External Merge, but also Internal Merge. In §4, I spell out some problems for the claim that music exhibits movement effects, and I provide an alternative analysis for the phenomena discussed by Katz & Pesetsky that does not allude to movement. I argue that this alternative account can equally well, if not better, explain the special behaviour of full cadences than the movement account does. §5 concludes.

2 Internal and External Merge in natural language

One of the highlights of the twenty-first-century developments in minimalism has been the operational unification of syntactic structure building and movement. While previous versions of minimalism (and its generative predecessors) took movement to involve a separate syntactic operation alongside Merge (or any other structure-building operation), Chomsky (2005) argued that nothing a priori forbids Merge to apply to previously created parts of the syntactic structure, and to remerge, or internally merge, these with the top node of the derivation (see also Starke 2001). Under this conception of Internal Merge, the question as to why natural language would display displacement operations no longer seemed to be in need of an explanation. If Merge is not restricted to External Merge, it would rather require additional explanation if language did not display movement effects.

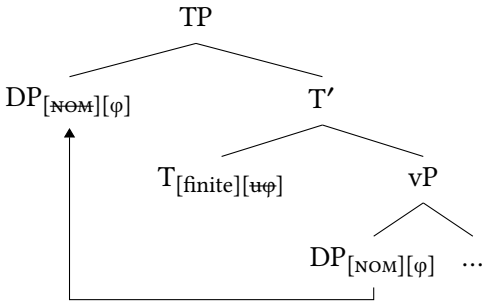
At the same time, questions still arise with respect to when Internal Merge should take place. Internal Merge differs from External Merge in the sense that Internal Merge does not have to take elements from the numeration into the

syntactic structure. If every element in the numeration needs to end up in the syntactic structure, it follows immediately that every element present in the numeration needs to undergo External Merge. But why would particular elements be required to undergo Internal Merge as well? From this perspective, there is no (external) reason that would force Internal Merge to take place.

The most straightforward solution would be to assume that Internal Merge only takes place if not applying it would render the sentence ungrammatical. Under that view, Internal Merge is a costly operation that only applies when necessary. This means that it is an operation for which a trigger is needed; and therefore, the question immediately arises as to what triggers Internal Merge.

Originally, it has been proposed by Chomsky (1995) that so-called uninterpretable features trigger movement. In a structure like (1), it is the uninterpretable [uφ] feature on T that triggers movement of the lower DP into the specifier position of the T-head, so that this feature, as well as the nominative feature on the DP, can be checked. The central conceptual motivation behind uninterpretable features as triggers for movement was that this would reduce two not well understood phenomena – the existence of semantically vacuous elements and the existence of displacement effects – to one not well understood notion: the need to remove uninterpretable features (where removal of uninterpretable features was said to take place under spec-head configuration).

(1)



This view, however, was later on rejected, primarily since it turned out that uninterpretable features could be checked at a distance (the uninterpretable feature probing down in its c-command domain to find a matching active goal). English expletive constructions (where the finite verb agrees with a lower VP-internal associated subject) (2), Icelandic quirky case constructions (where the verb agrees in number with a nominative object) (3), and various other constructions all underlie structures where the probe and the goal of agreement never appear in spec-head configuration:

(2) a. There seems to have arrived some student.
b. There seem to have arrived some students.

(3) Icelandic (Bobaljik 2008)
a. Jóni líkuđu thessir sokkar
Jon.DAT like.PL these socks.NOM
'Jon likes these socks.'
b. Mér virdast hestarnir vera seinir
me seem. PL the.horses be slow
'It seems to me that the horses are slow.'

If uninterpretable features can no longer be taken to trigger Internal Merge, the question arises as to what should do instead. Chomsky (2000; 2001) argues that movement should be thought of as an operation dependent on, and not triggered by, agreement. For him, probes, carrying uninterpretable features, could be equipped with an additional feature [EPP], which requires that the specifier of the probing head be filled. If no other suitable candidate could be merged externally in that position (such as an expletive subject like English *there*, or a dative subject, to the extent that such elements could be externally merged in this position in the first place; cf. Chomsky 2000; Deal 2009 for different proposals and discussion), the goal would raise into that position.

Even though using the EPP-feature gets these facts right, its postulation has often been criticized for a lack of independent motivation. The EPP-feature is rather a movement-triggering diacritic and does not build upon any explanation as to why movement should take place in the first place, although it could be that the presence or absence of movement (diacritics) is really just formal arbitrariness (a position taken by Biberauer et al. 2009; 2014; Biberauer & Roberts 2015, among others). For this reason, others have proposed to reinstall uninterpretable features themselves, rather than EPP-features, to be the sole triggers of movement (e.g., Bjorkman & Zeijlstra 2019). Nevertheless, whether uninterpretable features or subfeatures of uninterpretable features are the trigger for movement, in both cases uninterpretable features still form necessary elements in movement-triggering configurations.

Naturally, it is not the case that EPP-features and (un-)interpretable features are the only candidates for being movement triggers. Richards (2016), for instance, has argued that phonological adjacency requirements trigger movement; and Neeleman & Van de Koot (2008) have argued that movement may feed various mapping rules. But it should be noted that this type of approaches also relates the necessity of movement to interface requirements, as do uninterpretable

feature approaches. This all suggests that, in cognitive systems that lack formal features mediating between phonological and semantic features, triggering of Internal Merge might not be possible.

3 Internal and External Merge in music

In this section, I discuss the extent to which Merge can be said to be the (sole) structure-building operation in music, as claimed by Katz & Pesetsky. In order to provide evidence for this claim, Katz & Pesetsky build upon the insights presented in Lerdahl & Jackendoff's (1983) *Generative theory of tonal music* (GTTM). I will first briefly illustrate the major components of GTTM that are relevant for the discussion in this paper, without doing justice to the richness of this theoretical framework (§3.1). Then, in §3.2, I will present a particular aspect of music, namely the existence of structural hierarchies in music, which, for Katz & Pesetsky, forms evidence for their claim that musical structures are generated by at least External Merge. In §3.3, I discuss how, according to Katz & Pesetsky, other musical properties provide evidence for Internal Merge in music.

3.1 Lerdahl & Jackendoff's Generative theory of tonal music

According to the GTTM model, there are four components that determine the proper analysis of a musical structure. These four components are listed/given in (4) below:

(4) a. grouping structure
    b. metrical structure
    c. time-span reduction (TSR)
    d. prolongational reduction (PR)

Following Lerdahl & Jackendoff (1983: 8–9), grouping structure "expresses the hierarchical segmentation of the piece into motives, phrases, and sections"; metrical structure "expresses the intuition that the events of the piece are related to a regular alternation of strong and weak beats at a number of hierarchical levels"; TSR "assigns to the pitches of the piece a hierarchy of "structural importance" with respect to their position in grouping and metrical structure"; and PR, finally, "assigns to the pitches a hierarchy that expresses harmonic and melodic tension and relaxation, continuity and progression".

For Lerdahl & Jackendoff (1983), each component can assign a set of structures to a given string of music; and an additional set of preference interface rules then

determines which of these analyses is the correct one (often just one). In this sense, the musical architecture forms a strong resemblance with Jackendoff's parallel architecture of grammar (Jackendoff 1997; 2002; Culicover & Jackendoff 2005), which treats phonology, syntax, and semantics as independent generative components whose structures are also linked by interface rules: each component generates (a number of) structures, and interface rules determine what the proper mappings between these structures are. Such interface rules, for instance, determine which prosodic and which syntactic structures correlate.

Jackendoff's parallel architecture differs from Minimalist grammar in the sense that parallel architecture grammar has multiple engines, whereas Minimalist grammar has only one engine: its output leading to different levels of representation (phonetic form (PF) and logical form (LF)). However, at least according to Katz & Pesetsky, and I follow them in this respect, it is not the case that every musical component may bi-directionally inform every other component. Rather, it turns out that the outputs of grouping structure and metrical structure both inform TSR, which, in turn, informs PR. But if that is the case, the model for a grammar of music can be thought of as these components being directionally ordered, much like different grammatical components are directionally ordered in Minimalist grammar (Figure 3.2). Katz & Pesetsky's implementation of GTTM (Figure 3.1) is the reverse of the reverse Y-model.

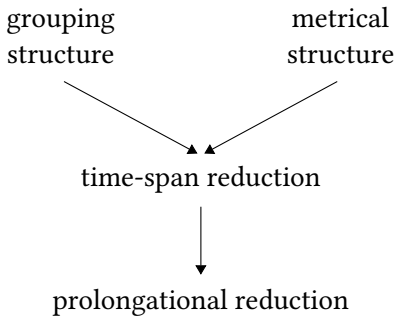


Figure 3.1: Katz & Pesetsky's (2011) (reverse reverse) Y-model of the grammar of music

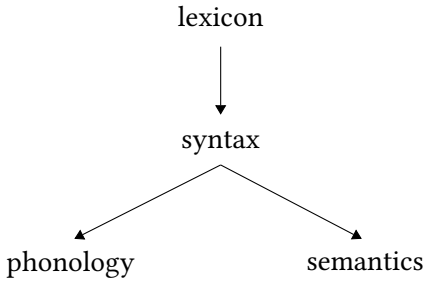


Figure 3.2: The reverse Y-model of the grammar of natural language

If this implementation is correct, the architecture of musical grammar forms a striking correspondence with the architecture of natural language grammar. A particular input is assigned an initial structure that can be derivationally transformed in subsequent structures, with particular well-formedness conditions holding at different levels of representation.

Under this architecture, it can indeed be investigated what the exact parallels are between the syntax of music and the syntax of natural language, and, most notably, whether the differences attested between language and music are merely a consequence of the differences in their building blocks or whether these differences are richer in nature.

3.2 External Merge in music

For Lerdahl & Jackendoff and for Katz & Pesetsky, the correspondence between language and music is stronger than merely being an architecture with various components that together are responsible for the analysis of a structure (irrespective of whether these components are derivationally or representationally connected by means of interface rules). As Lerdahl & Jackendoff already proposed, TSR in GTTM is very similar to prosodic structure in natural language, as both are formulated in terms of relative prominence. Moreover, Katz & Pesetsky take PR to align with linguistic syntax. The reason for them is that both PR and linguistic syntactic structures are binary branching, endocentric (i.e., headed) structures of the kind that is created by (External) Merge in Minimalist grammar. That such structures are headed can be witnessed by the fact that such structures are able to encode dependency relations between non-string-adjacent elements.

To see this, let us focus on the structure of PR. PR structures assign to the pitches a hierarchy that expresses harmonic and melodic tension and relaxation, continuity and progression. Simplifying things, every pitch that increases some kind of tension needs to be followed by some kind of relaxation. However, this need for tension followed up by relaxation is crucially not a string-adjacent condition. In fact, as we will see later on, it may very well be the case that the first tonic already induces a tension that is to be relieved by the final tonic, thus creating a constituent of two sisters whose heads span the entire musical piece. That means that tensions and relaxations in musical structures form non-local dependencies that are best explained as structural dependencies. This intuition is encoded in PRs by assigning head status to any sister of a node that is more relaxed. As an example, take the toy melody in Figure 3.3.

In this structure, the first event (the tonic C) establishes a sisterhood relation with the second event, the tonic being the head. In Western tonal music, tonics are always the most relaxed pitches, whereas pitches or chords based on pitches belonging to other scale degrees are felt to be tenser. Accordingly, the first event in this toy melody is the head of the merger with the second, third, fourth, and fifth events. The fifth event is the dominant (five degrees away from the tonic), which is tensed with respect to the tonic, but more relaxed with respect to the

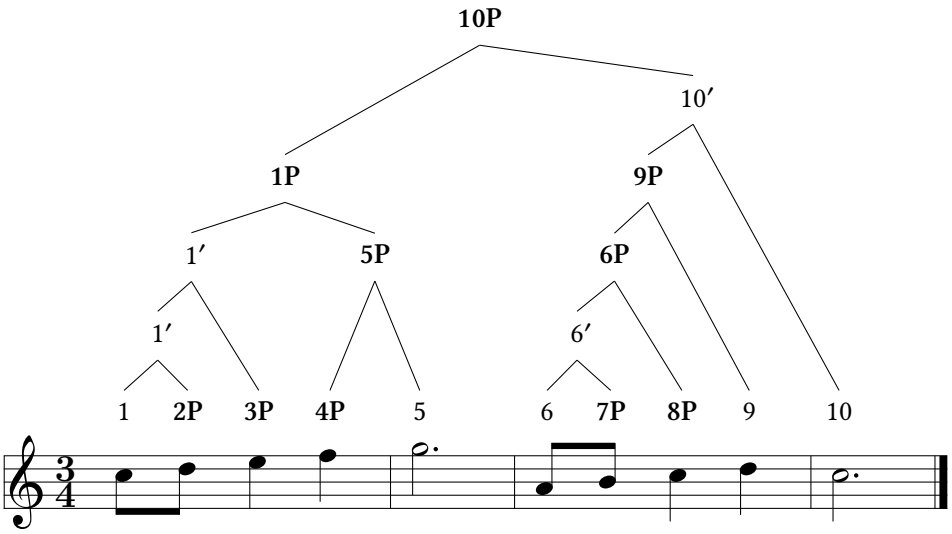


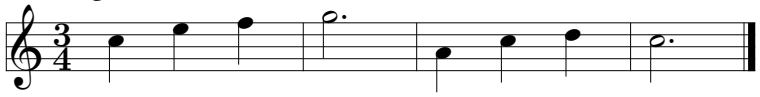
Figure 3.3: Toy melody (Katz & Pesetsky 2011: 16)

so-called subdominant (here, the fourth event), which is four degrees away from the tonic. Similarly, the final pitch (again, a tonic C) creates similar dependencies with the sixth till ninth events. The overall structure then consists of a constituent of two phrases: one in which the tonic in the first event is the head (1P) and one in which the tonic in the tenth event is the head (10P).

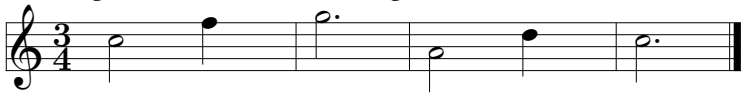
Evidence for this procedure of structure assignments comes from so-called *Schenkerian reductions* (see Forte 1959). Schenkerian reductions are best understood as musical summaries. Going bottom-up, removing every layer of non-heads will still yield a melody that feels like the same kind of melody as the intact structure. This process can in principle be continued until the most prominent chords are left. By contrast, if an event with higher prominence is left out, the piece is no longer perceived as a proper reduction. Examples, taken again from Katz & Pesetsky (2011), are presented below:

(5) Good reductions of Figure 3.3

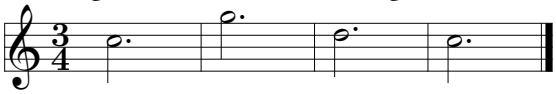
a. Deleting the non-heads of the lower 1′ and of 6′



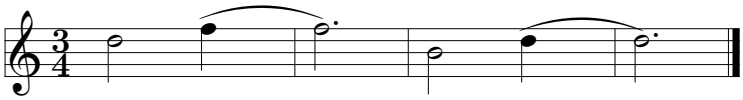
b. Deleting the non-heads of the higher 1′ and of 6P



c. Deleting the non-heads of the higher 1′, 5P and 9P



(6) A bad reduction of Figure 3.3



What does this tell us about Merge in music? The crucial comparison is that the structure-building operation appears to be similar to (External) Merge. Every two musical objects (being atomic or non-atomic) may merge and form a constituent of which the label is the same as that of one of its two daughters (the head). But if that is correct, it can be seen as evidence for there being a "syntactic engine" that is equally active in language and in music. This would, of course, be fully in line with Katz & Pesetsky's identity thesis for language and music. It is the module-specific properties of music that determine what elements can be merged and, once merged, which ones yield the heads (in terms of tension and relaxation, to be computed on the basis of scalar distance with respect to the tonic). But the combinatorial mechanism, Merge, applies to musical objects in exactly the same way as it applies to syntactic objects.

3.3 Internal Merge in music

The previous discussion of External Merge in music sets the ground for the next step in the discussion. If musical structures are indeed built by means of the single generative operation Merge (and the evidence for that claim, confirming the identity thesis for language and music, seems quite strong), then the question arises as to whether only External Merge applies or whether Internal Merge may apply as well. Formally, there is nothing in the combinatorial procedure that would exclude Internal Merge applying to music. Katz & Pesetsky argue that movement effects can indeed be attested in music. Let us first look at the arguments they present for that.

In order to assess whether musical pieces may display movement effects, one should first determine what the proper characteristics of movement in music would be. That task is far from trivial, as general diagnostics for movement (the surface position of some element does not correspond with the locus of its semantic interpretation) do not apply in music, for the simple reason that musical structures lack semantic interpretation (in the sense that musical structures

lack LF). Therefore, the diagnostics for movement should either be formal or PF-like. Moreover, such diagnostics are arguably different for phrasal movement and for head movement. Since Katz & Pesetsky do not provide any evidence for the existence of phrasal movement in music (even though they explicitly do not rule it out per se), but rather focus on head movement only, I will also only discuss what the characteristics of head movement in music would be. The characteristics that Katz & Pesetsky apply for head movement in language and music are given in (7) and (8), respectively:

(7) Head-movement in language (Katz & Pesetsky 2011: 40)
   a. Once the head H of a phrase HP has undergone head movement, H is pronounced string-adjacent to the head of a higher phrase, but at the same time …
   b. … the rest of HP remains an independent phrase that behaves just like a phrase whose head has not moved – even though:
   c. The movement is obligatory. Movement of finite V to T in French satisfies some need of an element in this structure […].
   d. The zero-level head that undergoes head movement to another zero-level head ends up tightly coupled to its new host. The two heads end up behaving like a single morphologically complex word for later processes of grammar (both syntactic and phonological).

(8) Head-movement in music (Katz & Pesetsky 2011: 41)
   a. Some chord X must be performed string-adjacent to a chord Y. But at the same time …
   b. … X has a normal set of syntactic dependents of its own, linearized normally – and thus apparently also heads its own phrase (an XP);
   c. The movement should be obligatory, insofar as it produces an alteration in the features of Y that is required in order for the derivation to succeed;
   d. Even though X may take a normal set of syntactic dependents, X is tightly coupled to its host Y, such that they function as an indivisible unit for other purposes (cf. the notion word).

Here, I will not contest these characteristics for movement, although I would like to point out that these characteristics should be interpreted in a uni-directional way. They are not diagnostics. Even if all effects attributed to head movement are indeed attested, this does not entail that the reverse must be the case

as well. If some α and β are both heads, pronounced string-adjacently, with α altering some feature of β and α and β together taken to form an indivisible unit (i.e., behaving word-like), this does not necessarily entail that α underwent head movement into β. I will come back to that in §4.

Katz & Pesetsky continue their argument by showing that so-called full cadences are a musical phenomenon that shows all the characteristics of head movement. In full cadences, the final chord, the tonic, which determines the key and counts as the head of the entire musical structure, must be preceded by a dominant, a chord whose root is five scale-steps away from the tonic and which has at least one dependent, generally headed by the so-called subdominant, often four scale-steps away from the tonic. In PR, the dominant is directly subordinate to the tonic and occupies a highly prominent position; metrically, it is often felt to be a much weaker chord that seems more deeply embedded in PR and seems to act as a weaker dependent of the tonic. This latter phenomenon is generally referred to as *cadential retention* – the phenomenon that the dominant and the tonic behave almost like a joint chord (and are even analysed as such in GTTM). An example is provided in Figure 3.4, where the dotted arrow (for now) indicates the stronger dependency of the dominant (δ) on the tonic (τ) (ν indicating the subdominant).

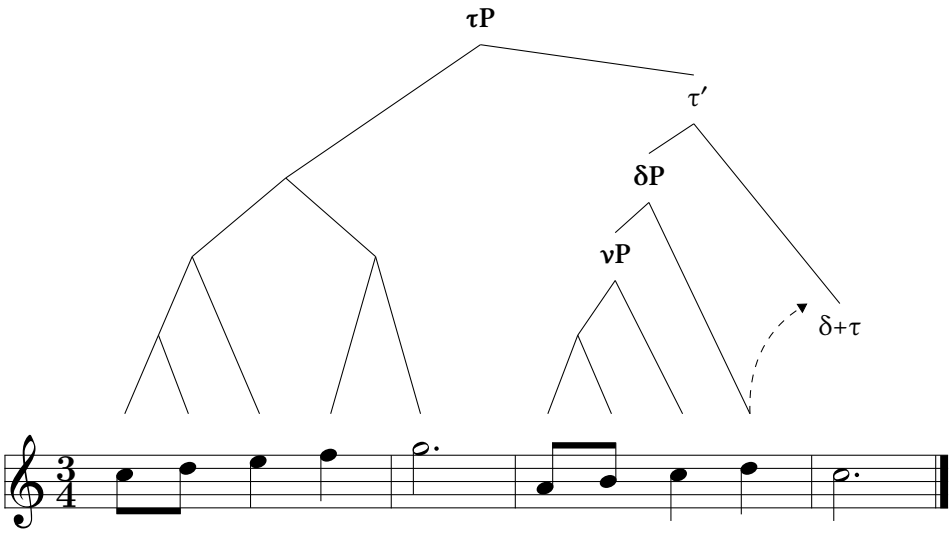


Figure 3.4: Example of a full cadence (Katz & Pesetsky 2011: 44)

Looking at the characteristics of head movement in music, Katz & Pesetsky conclude that full cadences indeed are the result of head movement, and, therefore, of the application of Internal Merge in music.

As for the first two characteristics, if the dominant indeed raises into the head position of the tonic (yielding the structure in (9), where angled brackets indicate lower copies of moved elements), the dominant is expressed string-adjacently to the tonic, even though the dominant still heads a phrase of its own (δP). This way, the construction behaves exactly like the first two clauses of the list of characteristics for head movement (in music).

(9) $[_{\tau P}\ [_{\delta P}\ [_{\nu P}\ \nu \ldots\ ]\ \langle\delta\rangle\ ]\ \delta{-}\tau]$

As for the third characteristic, Katz & Pesetsky claim that movement of the dominant into the tonic marks the tonic for establishing the key of the entire musical piece. They suggest that, in full cadences, movement of the dominant into the tonic head has the function of tonic-marking τ, i.e., assigning it the feature [+TON]. When the tonic head in a structure is tonic-marked, the terminal nodes of the phrase headed by the tonic are understood to belong to the key of τ. In this sense, head-movement of the dominant alters the tonic in having the feature [+TON].

As for the fourth characteristic, finally, Katz & Pesetsky argue that moving the dominant into the tonic position makes the joint dominant–tonic complex act more like a single unit in terms of metric position and makes the dominant look structurally less important than its PR position would legitimize. This joint behaviour, then, is what underlies the phenomenon of cadential retention.

On the basis of this analysis, Katz & Pesetsky conclude that musical structures are indeed generated by means of Merge, and the fact that Merge comprises both External and Internal Merge predicts that musical structures may indeed exhibit movement effects, of which full cadences are then an example. And, if musical structures indeed allow for movement, this forms additional evidence for Merge being the generator of musical structures. However, the reverse is not the case. If it turns out that head movement in music are absent (and that full cadences call for an alternative explanation), the claim that Merge is the sole generator of musical structures, and therefore also the identity thesis for language and music, can still be maintained. The evidence for structural (non-adjacent) dependencies in music and the structural mappings suffice as evidence for (External) Merge. The only question that would arise if (head) movement turns out to be absent in music, is: why is it absent in music despite the generative operation Merge being able to create structures involving movement, whereas (head) movement is so abundantly present in natural language? However, as argued for in §1 and §2, if so-called uninterpretable features are the sole triggers of Internal Merge and those features are absent in music, it is actually predicted that Internal Merge cannot apply in music.

4 Challenging movement in music

Full cadences are the sole cases of alleged (head) movement in music that Katz & Pesetsky present. That means that the validity of the claim that music exhibits movement rests solely on the validity of the argumentation behind their analysis of full cadences as involving head movement. Consequently, in order to maintain that Internal Merge applies in music, it must be shown that (i) full cadences indeed exhibit all the characteristics of head movement and (ii) that these constructions cannot be analysed in alternative terms (or that such an alternative analysis is much weaker). In this section, I argue that full cadences do not show a full parallel with instances of head movement in natural language and that the construction itself calls for an alternative analysis.

One fact that already casts doubt on the claim that music exhibits movement effects is that, outside full cadences, no other clear cases of movement in music have been attested. This is not because Katz & Pesetsky have been the first to look at those effects (although, admittedly, there have been few studies of the kind). Rohrmeier & Neuwirth (2014) discuss particular configurations that may involve movement in music as well, but crucially state that these constructions do not have to be analysed as syntactic movement and therefore do not form any evidence in favour of movement in music. The only other claim of movement in music that I am aware of is Temperley (1999), who notes a parallel between syncopation in rock music and head movement in syntax.

Strikingly, these cases of alleged movement in music are the linguistic equivalent of rightward, string-adjacent head-movement. That, of course, already triggers the question as to why other instances of movement (phrasal movement, non-string-adjacent movement and leftward movement) have so far not been attested in music.

It should be noted in this respect that the core cases of movement in language indeed are cases of leftward, non-string-adjacent movement. That phrasal movement has not been attested as such is not so telling. Both head movement and phrasal movement are indeed solid cases of movement, although head-movement has often been said to be an instance of PF-movement, instead of movement that takes place in narrow syntax (cf. Chomsky 1995; Boeckx & Stjepanović 2001; Harley 2004). However, even if head movement were an instance of PF-movement, this would not invalidate the claim that music exhibits movement effects, as musical structures, just like syntactic structures in language, are to be linearized. In fact, one might even argue that the specific nature of music (with its sole sound side and lack of a meaning side) would rather call for head movement only.

Things are different, however, when it comes to rightward, string-adjacent movement, which has received more scepticism in the linguistic literature. Rightward movement, especially in comparison to leftward movement, is heavily constrained (cf. Ross 1967; Kayne 1994; Cinque 1996; Ackema & Neeleman 2002; Abels & Neeleman 2012). For instance, Kayne (1994) observes that there are verb-second languages but no so-called verb-penultimate languages (where the finite verb appears in the penultimate position). Neither are there languages where *Wh*-terms consequently move to the right (with the possible exception of certain sign languages, cf. Cecchetto et al. 2009). According to Abels & Neeleman (2012), rightward phrasal movement is only possible for full extended projections (that do not strand any parts of it), and according to Ackema & Neeleman (2002), rightward head movement is restricted to moving heads that do not cross any of their dependents. If that is correct, then rightward head movement can only be string-adjacent.

But string-adjacent movement perhaps even calls for more scepticism. How can one determine whether a particular element underwent movement if the linear position of the moved element is the same as its base position? Already in linguistics this is far from clear. In the case of string-adjacent phrasal movement, there might be good reasons to assume that some particular elements indeed undergo movement. For instance, Pesetsky (1987) and Bobaljik (1995; 2002) have argued that subject *Wh*-phrases (like *Who* in *Who left?*) arguably undergo movement from Spec,TP into Spec,CP (to end up in A-bar position) (pace Grimshaw 1997). For head movement things are less clear. Do heads in head-final languages (the only candidates for rightward string-adjacent head movement), such as Korean and Japanese, undergo head movement or not? Is it the case that, in such languages in a configuration like (10), V moves into T and/or T into C?

(10) $[_{CP}\ [_{TP}\ [_{VP}\ V\ ]\ T\ ]\ C\ ]$

Whether languages like Japanese and Korean exhibit string-adjacent rightward head movement or not has been widely discussed in the literature. Various scholars have provided arguments in favour of it. Otani & Whitman (1991) have argued that, in Japanese, the verb must raise to account for various ellipsis effects. The same applies to Koizumi (1995; 2000), who has primarily discussed scrambling and coordination. Also, Yoon (1994) makes an argument in favour of string-adjacent head movement based on coordination of tensed and untensed conjuncts. Choi (1999), finally, formulates an account in terms of NPI licensing that calls for string-adjacent head movement. But as Han et al. (2007; 2016) have shown, basing themselves on arguments by Kim (1995), Chung & Park (1997), Hoji

(1998), Kim (1999), and Fukui & Sakai (2003), all these facts can also be accounted for by approaches that do not allude to rightward head movement. In turn, Han et al. (2007; 2016) argue that head-final languages (Korean is their example) may actually vary language-internally with respect to whether heads undergo raising or not (though see Zeijlstra 2017 for an argument against their claim that some varieties of Korean provide evidence for string-adjacent head movement).

But even if in some languages string-adjacent, rightward head movement can be attested, this does not predict that this is the case for every language. There may be particular language-specific reasons that call for such instances of string-adjacent, rightward head movement, but that does not entail that, in every head-final language, verbs raise into higher heads of the extended projection.

Under the null hypothesis that one should only postulate movement to take place if the data cannot be accounted for otherwise, the question really arises how strong the evidence for movement of the dominant into the tonic position is. What would go wrong if one were to analyse full cadences as instances where the dominant does not raise into the tonic-position but instead just stays in its string-adjacent PR position?

For this, we need to reinvestigate the characteristics of full cadences presented in §3.3. It turns out that, out of the four listed properties, three of them immediately follow by assuming that the dominant stays in situ (11). The fact that the dominant is expressed string-adjacently to the tonic, and the fact that the dominant still heads a phrase of its own (δP) are fully compatible with the analysis in (11).

(11) $[_{\tau P}\ [_{\delta P}\ [_{\nu P}\ \nu \ldots\ ]\ \delta\ ]\ \tau]$

Moreover, the fact that the dominant and the tonic are perceived as one unit (the musical counterpart of being a single word) can also be explained under string-adjacency. Here, the parallel with affixation comes up. Under more traditional concepts of head movement heads raise into higher head positions to ensure realization of the higher head as an affix on the lower head (or vice versa). In that sense, head movement is triggered by the so-called *stray-affix filter* (cf. Lasnik 1981; 1995; Baker 1988) (in any of its guises). For this stray-affix filter to apply, it suffices that the two relevant heads always appear in a string-adjacent position at PF. Now, in head-initial languages, this cannot be guaranteed without alluding to verb movement (due to intervening specifiers/adjuncts), but in head-final languages, where heads are already string-adjacent to each other, it can. Following Bobaljik (1995), an affix can be spelled out on the verb in an OV-language without the verb moving to it, since V and the affix are string-adjacent

at PF. But if that is the case, string-adjacency can suffice as a condition for the dominant and the tonic to be realized as a single unit. Consequently, the fact that the dominant and the tonic end up as one unit does not form evidence for head movement.

This leaves the obligatoriness of head movement as a final possible piece of evidence in favour of an analysis of full cadences in terms of head movement. Head movement in language is obligatory (e.g., movement of finite V to T in French must take place; the finite verb cannot stay in situ). This obligation for head movement is generally understood as a movement-triggering requirement: Some feature of the higher head must be altered for the derivation to proceed, and only raising of another head into this position can establish this feature alteration. For movement, Katz & Pesetsky argue that this feature alteration must be understood as tonic-marking. Movement of the dominant into the tonic position assigns a feature [+TON] to the tonic. Having a tonic feature, in turn, is responsible for this tonic to establish the key of the entire musical piece.

Two questions come to mind here. First, is it necessary that movement triggers such a feature alteration? Can't adjacency suffice here as well? It is known from various impoverishment facts that features present on one head can manipulate the features on a neighbouring head without undergoing movement. Hence, even if the tonic must be tonic-marked by the dominant, this does not have to be realized by means of movement.

Second, is it really the case that the feature of the tonic must be tonic-marked? After all, full cadences are not obligatory in music. Tonics do not require dominants to remerge into their head positions, and neither is it impossible for a dominant to remain in situ (which generally appears to be the case, except perhaps for full cadences). In that sense, head movement of the kind in music is not obligatory in the sense we understand movement to be obligatory in language. What appears to be the case under Katz & Pesetsky's analysis is that movement of the dominant into the tonic is only obligatory under string-adjacency, a much weaker requirement.

But if the structure underlying full cadences is not obligatory for tonic-marking, what one can say is that, at best, it facilitates key establishment. It may help the listener in determining what the key of the entire phrase or piece is. But naturally, other musical facts may play a similar role. For instance, the selection of pitches used in the musical piece already forms a strong (and often sufficient) cue for establishing the key of the entire piece. And also, if harmonic properties determine the PR of a musical piece and if TSR–PR mismatches may only take place under particular circumstances that follow from the underlying PR structure, such mismatches may also provide the listener with a cue of what the key

of the entire piece is. In other words, what full cadences seem to do is facilitate key recognition instead of establishing it.

This all calls for an alternative picture for an analysis of full cadences along the lines of (11), where the adjacency of the dominant and the tonic results in a confirmation of the tonic determining the key and where cadential retention is nothing but the result of an adjacency requirement (a string-adjacent dominant and tonic may or must be realized as a single unit). Already the existence of a viable alternative to the head-movement analysis undermines the status of full cadences as evidence for head movement in music. And this alternative analysis may equally well get the facts right, if not better. But if the only piece of evidence in favour of movement in music turns out to be inconclusive (and may be even incorrect), there is no evidence left any more for the claim that music triggers Internal Merge.

So where do we stand? If full cadences can be equally well, if not better, understood in terms of adjacency requirements, much like Bobaljik (1995) takes such requirements to suffice to establish dependencies between adjacent heads at PF, there appears to be no evidence for movement in music. This allows us to entertain a stronger and more powerful hypothesis, namely that musical structures, despite being generated by Merge, do not exhibit any kind of movement. There is only External Merge going on in music. That amounts to saying that, despite the principled availability of its application, Internal Merge never takes place in music. Given the discussion in §1, where I have argued that that musical building blocks crucially lack the type of features that may trigger Internal Merge and that, consequently, the identity thesis for language and music should predict that Internal Merge never takes place in music, I take this to be a welcome result.

5 Conclusions

In this paper, I have aimed at rethinking remerge. Starting from the premise that uninterpretable features are the sole trigger of Internal Merge, I have looked at another cognitive system, music, to see whether in such a system, where, clearly, (un)interpretable features are absent, Internal Merge may still apply. Focussing on Katz & Pesetsky's elaboration and modification of Lerdahl & Jackendoff's (1983) Generative theory of tonal music, I have evaluated Katz & Pesetsky's claim that musical structures also exhibit movement, and, in particular, their claim that full cadences are to be understood as involving string-adjacent, rightward head movement. My conclusion is that full cadences are equally well, if not better, understood in terms of linear adjacency requirements and that, therefore, the

presented evidence of movement in music does not hold. I have argued that this rather calls for a view of music where movement is absent. However, I have argued as well that this does not speak against Katz & Pesetsky's identity thesis for language and music, but rather speaks in favour of it. Musical structures indeed appear to be generated by means of Merge. However, the absence of uninterpretable features in music prevents Internal Merge from applying in the first place, at least under the assumption that uninterpretable features are the sole trigger for the application of Internal Merge. The reason why music lacks (un)interpretable features is that (un)interpretable features can only emerge in cognitive systems whose building blocks are multi-modular, such as linguistic building blocks. Musical building blocks, by contrast, are mono-modular and can therefore never consist of such (un)interpretable features. The absence of movement in music thus follows directly from the differences between musical and linguistic building blocks and is, therefore, fully in line with Katz & Pesetsky's identity thesis for language and music.

Abbreviations

EPP	extended projection principle	NOM	nominative
GTTM	Generative theory of tonal music	PF	phonetic form
		PR	prolongational reduction
LF	logical form	TSR	time-span reduction

Acknowledgements

I am much indebted to David Pesetsky, who triggered my interest in this topic during his 2009 class on language and music at the EGG summer school in Poznan, and during subsequent discussions. I am also very grateful to Jonah Katz for providing very helpful comments on earlier versions of this work. This paper is the outcome of one of my "Hot topics in language and cognition" classes, taught at the Cognitive Science Center Amsterdam and at the University of Göttingen in 2012–2013, where I discussed Katz & Pesetsky's paper. I thank my students for valuable feedback. Previous versions of this paper have been presented at GLOW in Asia X (held at the National Tsing Hua University, Taiwan), WCCFL 35 (held at Simon Frasier University, Vancouver) and at the Workshop "What drives syntactic computation? Alternatives to formal features", which was part of the annual DGfS meeting in Leipzig in 2015. I would like to thank the organizers and audiences of these events for the opportunity and their comments. All errors, of course, are mine.

References

Abels, Klaus & Ad Neeleman. 2012. Linear asymmetries and the LCA. *Syntax* 15(1). 25–74. DOI: 10.1111/j.1467-9612.2011.00163.x.

Ackema, Peter & Ad Neeleman. 2002. Effects of short-term storage in processing rightward movement. In Sieb Noteboom, Fred Weerman & Frank Wijnen (eds.), *Storage and computation in the language faculty*, 219–256. Dordrect: Kluwer.

Baker, Mark C. 1988. *Incorporation: A theory of grammatical function changing*. Chicago: University of Chicago Press.

Biberauer, Theresa, Anders Holmberg & Ian Roberts. 2009. Linearization and the architecture of grammar: A view from the Final-over-Final Constraint. In Vincenzo Moscati & Emilio Servidio (eds.), *Proceedings of XXXV Incontro di Grammatica Generativa*, vol. 3 (University of Siena CISCL Working Papers), 77–89. MITWPL.

Biberauer, Theresa, Anders Holmberg, Ian Roberts & Michelle Sheehan. 2014. Complexity in comparative syntax: The view from modern parametric theory. In Frederick J. Newmeyer & Laurel B. Preston (eds.), *Measuring grammatical complexity*, 103–127. Oxford: Oxford University Press.

Biberauer, Theresa & Ian Roberts. 2015. Clausal hierarchies. In Ur Shlonsky (ed.), *Beyond functional sequence* (The cartography of syntactic structures 10), 295–313. Oxford: Oxford University Press.

Bjorkman, Bronwyn M. & Hedde Zeijlstra. 2019. Checking up on (φ-)Agree. *Linguistic Inquiry* 50(3). 527–569. DOI: 10.1162/ling_a_00319.

Bobaljik, Jonathan David. 1995. *Morphosyntax: The syntax of verbal inflection*. MIT. (Doctoral dissertation).

Bobaljik, Jonathan David. 2002. A-chains at the PF-interface: Copies and "covert" movement. *Natural Language & Linguistic Theory* 2(20). 197–267. DOI: 10.1023/A:1015059006439.

Bobaljik, Jonathan David. 2008. Where's phi? Agreement as a postsyntactic operation. In Daniel Harbour, David Adger & Susana Béjar (eds.), *Phi theory: Phi-features across modules and interfaces*, 295–328. Oxford: Oxford University Press.

Boeckx, Cedric & Sandra Stjepanović. 2001. Head–ing toward PF. *Linguistic Inquiry* 32(2). 345–355. DOI: 10.1162/00243890152001799.

Bošković, Željko. 2007. On the locality and motivation of Move and Agree: An even more minimal theory. *Linguistic Inquiry* 38. 589–644. DOI: 10.1162/ling.2007.38.4.589.

Brody, Michael. 1997. Perfect chains. In Liliane Haegeman (ed.), *Elements of grammar*, 139–167. Dordrecht: Kluwer.

Cecchetto, Carlo, Carlo Geraci & Sandro Zucchi. 2009. Another way to mark syntactic dependencies: The case for right-peripheral specifiers in sign languages. *Language* 85(2). 278–320. DOI: 10.1353/lan.0.0114.

Choi, Young-Sik. 1999. Negation, its scope and NPI licensing in Korean. In Rebecca Daly & Anastasia Riehl (eds.), *ESCOL [the Eastern states conference on linguistics] '99*, 25–36. Ithaca, NY: CLC Publications.

Chomsky, Noam. 1995. *The Minimalist program.* Cambridge, MA: MIT Press.

Chomsky, Noam. 2000. Minimalist inquiries: The framework. In Roger Martin, David Michaels & Juan Uriagereka (eds.), *Step by step: Essays on minimalist syntax in honor of Howard Lasnik*, 89–155. Cambridge, MA: MIT Press.

Chomsky, Noam. 2001. Derivation by phase. In Michael Kenstowicz (ed.), *Ken Hale: A life in language*, 1–52. Cambridge, MA: MIT Press.

Chomsky, Noam. 2005. Three factors in language design. *Linguistic Inquiry* 36(1). 1–22. DOI: 10.1162/0024389052993655.

Chung, Daeho & Hong-Keun Park. 1997. NPIs outside of negation scope. In Homin Sohn & John Haig (eds.), *Japanese/Korean linguistics*, vol. 6, 415–435. Stanford, CA: CSLI Publications.

Cinque, Guglielmo. 1996. The antisymmetric programme: Theoretical and typological implications. *Journal of Linguistcs* 32(2). 447–464. DOI: 10.1017/S0022226700015966.

Culicover, Peter W. & Ray Jackendoff. 2005. *Simpler syntax.* Oxford: Oxford University Press.

Deal, Amy Rose. 2009. The origin and content of expletives: Evidence from "selection". *Syntax* 12(4). 285–323. DOI: 10.1111/j.1467-9612.2009.00127.x.

Forte, Allen. 1959. Schenker's conception of musical structure. *Journal of Music Theory* 3(1). 1–30. DOI: 10.2307/842996.

Fukui, Naoki & Hiromu Sakai. 2003. The visibility guideline for functional categories: Verb-raising in Japanese and related issues. *Lingua* 113(4–6). 321–375. DOI: 10.1016/S0024-3841(02)00080-3.

Grimshaw, Jane. 1997. Projection, heads, and optimality. *Linguistic Inquiry* 28(3). 373–422.

Han, Chung-hye, Jeffrey Lidz & Julien Musolino. 2007. V-raising and grammar competition in Korean: Evidence from negation and quantifier scope. *Linguistic Inquiry* 38(1). 1–47. DOI: 10.1162/ling.2007.38.1.1.

Han, Chung-hye, Julien Musolino & Jeffrey Lidz. 2016. Endogenous sources of variation in language acquisition. *PNAS* 113(4). 942–947. DOI: 10.1073/pnas.1517094113.

Harley, Heidi. 2004. Merge, conflation, and head movement: The First Sister Principle revisited. In Keir Moulton & Matthew Wolf (eds.), *NELS 34: Proocedings of the thirty-fourth annual meeting of the North East Linguistic Society*, 239–254. Amherst, MA: GLSA Publications.

Hoji, Hajime. 1998. Null object and sloppy identity in Japanese. *Linguistic Inquiry* 29(1). 127–152.

Jackendoff, Ray. 1997. *The architecture of language*. Cambridge, MA: MIT Press.

Jackendoff, Ray. 2002. *Foundations of language*. Oxford: Oxford University Press.

Katz, Jonah & David Pesetsky. 2011. The identity thesis for language and music. lingBuzz/000959.

Kayne, Richard S. 1994. *The antisymmetry of syntax*. Cambridge, MA: MIT Press.

Kim, Jong-Bok. 1995. On the existence of NegP in Korean. In Susumo Kuno (ed.), *Harvard studies in Korean linguistics: VI. Proceedings of the 1995 Harvard International Symposium on Korean, January 13–15, 1995*, 267–282. Cambridge, MA: Harvard University.

Kim, Soowon. 1999. Sloppy/strict identity, empty objects, and NP ellipsis. *Journal of East Asian Linguistics* 8(4). 255–284. DOI: 10.1023/A:1008354600813.

Koizumi, Masatoshi. 1995. *Phrase structure in minimalist syntax*. Cambridge, MA: MIT. (Doctoral dissertation).

Koizumi, Masatoshi. 2000. String vacuous overt verb raising. *Journal of East Asian Linguistics* 9(3). 227–285. DOI: 10.1023/A:1008311420647.

Lasnik, Howard. 1981. Restricting the theory of transformations: A case study. In Norbert Hornstein & David W. Lightfoot (eds.), *Explanation in linguistics*, 152–173. London: Longman.

Lasnik, Howard. 1995. Verbal morphology: Syntactic structures meets the minimalist program. In Hector Campos & Paula Kempchinsky (eds.), *Evolution and revolution in linguistic theory*, 251–275. Georgetown: Georgetown University Press.

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

Neeleman, Ad & Hans Van de Koot. 2008. Dutch scrambling and the nature of discourse templates. *Journal of Comparative Germanic Linguistics* 11(2). 137–189. DOI: 10.1007/s10828-008-9018-0.

Otani, Kazuyo & John Whitman. 1991. V-raising and VP-ellipsis. *Linguistic Inquiry* 22(2). 345–358.

Pesetsky, David. 1987. Wh-in-situ. In Eric J. Reuland & Alice G. B. ter Meulen (eds.), *Movement and unselective binding*, 98–129. Cambridge, Massachusetts: MIT Press.

Richards, Norvin. 2016. *Contiguity theory*. Cambridge, MA: MIT Press.

Rohrmeier, Martin & Markus Neuwirth. 2014. Towards a syntax of the classical cadence. In Markus Neuwirth & Pieter Bergé (eds.), *What is a cadence? Theoretical and analytical perspectives on cadences in the classical repertoire*, 285–336. Leuven: Leuven University Press.

Ross, John Robert. 1967. *Constraints on variables in syntax*. MIT. (Doctoral dissertation).

Schlenker, Philippe. 2016. Prolegomena to music semantics. lingBuzz/002925.

Starke, Michal. 2001. *Move dissolves into Merge: A theory of locality*. University of Geneva. (Doctoral dissertation).

Svenonius, Peter. 2007. Interpreting uninterpretable features. *Linguistic Analysis* 33(3–4). 375–413.

Temperley, David. 1999. Syncopation in rock: A perceptual perspective. *Popular Music* 18. 19–40. DOI: 10.1017/S0261143000008710.

Yoon, James. 1994. Korean verbal inflection and checking theory. In Heidi Harley & Colin Phillips (eds.), *The morphology-syntax connection*, vol. 22 (MIT Working Papers in Linguistics), 251–270. Cambridge, MA: MITWPL.

Zeijlstra, Hedde. 2008. On the syntactic flexibility of formal features. In Theresa Biberauer (ed.), *The limits of syntactic variation*, 143–174. Amsterdam: John Benjamins. DOI: 10.1075/la.132.06zei.

Zeijlstra, Hedde. 2012. There is only one way to agree. *The Linguistic Review* 29(3). 491–539. DOI: 10.1515/tlr-2012-0017.

Zeijlstra, Hedde. 2017. *Two varieties of Korean or the ban against rightward head movement*. Paper presented at NELS [North East Linguistics Society] 48, University of Iceland, Reykjavik, 27–29 October, 2017.