No movement in music: a reply to Katz & Pesetsky (2011)

I. In an influential paper Katz & Pesetsky (2011) present the identity thesis for language and music, stating that 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. 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); particular musical structures can map into other musical structures (according to Lerdahl & Jackendoff 1983, time-span reductions map into prolongational reductions), just like syntactic structure maps into prosodic structure (cf. Selkirk 1984); and, finally, they take movement phenomena (i.e. the application of Internal Merge) to be present in both language and music. While fully endorsing the view that musical structures are the result of multiple application 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 we take it to speak in its favour: the identity thesis for language and music reduces all differences between language and music to its basic building blocks. If Internal Merge in language is driven by particular features (in casu uninterpretable features, cf. Chomsky 1995, 2002, Boskovic 2007) that are language-specific and not applicable to music, the direct consequence is that Internal Merge cannot apply in music either. II. The evidence Katz & Pesetsky provide in favour of movement in musical structure comes from the phenomenon full cadence (listen here for an example). In full cadences, the final chord, the tonic (τ), which determines the key and counts as the head of the entire musical structure, must be preceded by a dominant (δ), a chord whose root is five scale-steps away from the tonic, and which has at least one dependent, headed by the so-called subdominant (ν), generally, but not always four scale-steps away from τ. Whereas in terms of time-span reduction ν is a dependent of the head δ, in the prolongational reduction δ acts as a dependent of τ. δ and τ are felt to yield some unit. See (1) for a relevant excerpt, taken from Katz & Pesetsky. Since the structure requires some full δP as the complement of τ and at the same time δ and τ form a single unit, where δ acts as the dependent of τ, Katz & Pesetsky argue that δ undergoes head-movement to τ. III. These facts, however, do not necessarily reflect head movement. To see this, let’s think of δ as the musical counterpart of some affix that heads a projection of its own (δP) and at the same time must adjoin to some higher head (τ) in compliance with the Stray Affix filter (Lasnik 1981, 1995, Baker 1988). In head-initial configurations this naturally triggers head movement, as shown in (2a):

(2) a. [τP [δP δ-] τ]
   b. [τP [δP δ-] t₁]  
But in head-final configurations (2b), as Bobaljik (1995) has pointed out, head movement is not necessary at all, since the requirement that δ and τ are string adjacent, a necessary condition for the Stray Affix filter, is already fulfilled, thus rendering δ-to-τ movement superfluous. Since full cadences necessarily involve head-final tonics, all instances of full cadences can thus be explained in this way without alluding to head movement: full cadence simply requires that the δ and τ heads must be string-adjacent. The fact that full cadences are not necessarily indicative of head movement does not, however, entail that they do not involve head movement. The facts presented by Katz & Pesetsky are still fully compatible with a head movement analysis. But, there are three arguments that suggest that an analysis of full cadence in music that does not treat it in terms of head movement fares much better. First, full cadences one of the very few instances of alleged movement in music. Virtually no other structural musical phenomenon has been observed thus which must result from movement, suggesting that at best movement is extremely limited if not absent in music at all (Temperley 1999 relates that syncopation in rock music to displacement, but again, this would be an instance of string-adjacent movement, so the same criticisms apply here, and Rohrmeier and Neuwirth 2014 discuss a few more examples, but provide alternative analyses in terms of non-movement for all of them as well). If movement is altogether absence (and full cadence results from some...
adjacency requirement), this asymmetry is explained; if it is not, this asymmetry between language and music remains mysterious. Second, the kind of head movement that Katz & Pesetsky propose is the musical variant of rightward and string-adjacent movement. These are exactly the two types of movement that receive general scepticism in linguistic theory. See, for instance, Kayne 1995, Ackema & Neeleman 2002 for arguments against rightward movement – the arguments from the latter are stated in general cognitive terms, and would naturally apply to music as well; Kayne’s arguments are more language-specific, but could also be phrased in domain-general terms (e.g. as Biberauer et al t.a. do). See Bobaljik 1995 and many others working in the Distributed Morphology framework for arguments against string-adjacent head movement (arguing that head movement that is needed to create adjacency requirements need not take place if those requirements are already fulfilled, an argument that naturally extends to string-adjacent movement in musical). Third, if Internal Merge never takes place in music, this fact can very naturally be explained as a difference between musical and linguistic building blocks. The major difference between linguistic building blocks and musical building blocks is that linguistic building blocks are triplets of formal, semantic and phonological features, whereas musical building blocks are chords, consisting of multiple pitch-class elements (at least for Katz & Pesetsky). Clearly, the latter do not combine features that belong to different modalities (as linguistic building blocks do). Chords are mono-modular in this sense. A general property of uninterpretable features is that they are interface features (after the definition of Svenonius 2007). In order for some linguistic feature to be uninterpretable, it must be visible to two modules, but only give rise to well-formedness effects on of them. Phi-features, for instance, are interface features that are always formally/syntactically active, and at the same time semantically interpretable on only some elements (DPs). On other elements (e.g. finite verbs) they are uninterpretable and give rise to movement effects to check them off. If uninterpretable features must be interface features, the consequence is that cognitive building blocks, such as musical chords, that are monomodular can never be said to carry any uninterpretable features (by definition). Consequently, if Internal Merge is indeed driven by the checking requirements of uninterpretable features, Internal Merge in music must be categorically ruled out. IV. The view pertained here has two further consequences: (i) all instances of linguistic movement are triggered by uninterpretable features that have an interpretable counterpart (and not by [EPP] or sui generis uninterpretable features); and (ii) there is no trigger for movement in music. Both are, I argue, essentially correct. As for (i), I follow Bošković 2007 and Zeijlstra 2012 who argue that if Agree is taken to be strictly unidirectional all instances of movement must be driven by uninterpretable feature that have an interpretable counterpart. As for (ii), this is argued against by Katz & Pesetsky, who argue that the selection of key domains (for them a separate Tonal-Harmonic component, much alike to semantic component in language) interprets structures that must have undergone head movimiento. Head movement in music is driven by a featural need of the tonic: δ-to-τ movement allows the terminal nodes of a particular subtree of τP to be understood as belonging to the key of τ. To the extent that determination of key domains takes place ‘post-syntactically’, i.e. after the formation of musical structures (see Rohrmeier 2008, 2011 for a different view), the structures that form the input for key selection do not have to reflect movement effects. Again, since tonics and dominants are always head-final, every effect of alleged head-movement in music can be captured in terms of affixation under adjacency as well. If head movement of a dominant into a tonic may result into tonic marking, so may affixation of a dominant into a tonic (as long as these affixation effects are visible at some interpretative level). In fact, one might already even argue for the latter already on linguistic grounds, as head movement in language is generally taken to lack any interpretative effects. Hence, even if Katz & Pesetsky are correct in post-syntactic key assignment, there is still no evidence that this requires a mechanism that depends on movement and not on adjacency effects. Consequently, key assignment cannot be taken to be an argument in favour of movement in music.