Linearization and Local Dislocation: Derivational Mechanics and Interactions*

David Embick
University of Pennsylvania

1. Introduction

On the assumption that the syntax generates hierarchical representations that are accessed by both sound and meaning systems, it is an empirical question how the syntactic representation and the representation(s) referred to in different morphophonological processes relate to one another. In Chomsky and Halle [7:9], for example, this question is posed in terms of how two conceptions of *surface structure*, "output of the syntactic component" and "input to the phonological component," relate to one another, with identity being a possibility that is excluded because of the existence of cases in which these two notions appear to differ. What is then required is a theory of the possible relationships between syntactic and phonological structures, on the assumption that in spite of some differences, the overall patterns are systema tic. In terms of current models of syntax and its interfaces, this amounts to giving a theory of PF.

Since this set of questions was initially formulated, research in this area has identified a range of cases in which syntactic structure and phonological structure do not line up with one another, in a number of domains (prosodic phonology, cliticization, bracketing paradoxes, etc.). To the extent that phenomena of this type require syntactic and phonological representations that are distinct from one another, the further question is how great the differences are. I take it that the possible deviations are highly restricted in their scope, something that amounts to assuming a "restrained" view of PF. Within the context of a derivational framework, the program is to specify the different

* A version of this paper was presented at the colloquium at the University of Massachusetts, Amherst, the III Encuentro de Gramática Generativa at the University of Comahue in Neuquén, Argentina, and the colloquium of the CUNY Graduate Center. I am grateful to these audiences for raising a number of significant points that I have attempted to integrate into this written version. For additional comments on the work presented here I am indebted to Rajesh Bhatt, Morris Halle, Alec Marantz, Marjorie Pak, and Dominique Sportiche; I also would like to thank an anonymous reviewer for this edition of *Linguistic Analysis* for a number of helpful points.

Linguistic Analysis, 33 3-4 (2003)

P.O. Box 2418, Vashon, WA 98070

computations that augment and alter the syntactic representation. The central concern is thus to provide a theory in which sound/syntax connections (and thus sound/meaning connections) are as systematic as possible given the range of data to be accounted for.

In terms of specific proposals, one way of viewing a certain part of the research in the framework of Distributed Morphology is as an attempt to identify some of the relevant PF-mechanisms, and to answer attendant questions concerning their ordering, interaction, and so on. One component of this syntactic approach to morphology is a theory of the operations that apply on the PF branch of the grammar, with some traditionally "morphological" phenomena (allomorphy, phonological versus syntactic bracketing, syncretism) being addressed as part of a larger set of questions whose primary concern is the interface that mediates between hierarchical representations and their ultimate phonetic expression.

This article contributes to this line of research by examining the representation of linear order in the PF component of the grammar, as revealed by cases in which PF-rules affix one element to another under linear adjacency: *Local Dislocation* (LD). Part of what it means to have a theory of PF is to have a constrained theory of mismatches between syntax and morphophonology, and affixation under adjacency results in mismatches that any syntactic theory must account for. Moreover, by looking at the "special" cases—i.e., the cases in which some PF requirement triggers an operation resulting in a mismatch—some insight can be gained on linear representations in the normal case.

While the focus of this article is on the details of LD, it is important to view this operation in the context of a general theory of syntactic affixation. The fact that the conditions on affixation are sensitive to syntactic and post-syntactic notions of locality is a significant point with important architectural consequences; see Embick [14] and Embick and Marantz [18] for discussion. There are two primary components to the discussion below, which follow on some initial points about LD in section 2. Concentrating on the formal properties of LD, I examine in section 3 two different properties of linear representations that restrain the application of LD. It is argued that linearization statements that concatenate elements—and hence LD operations defined in terms of these—are *typed*, with the types being distinct structural objects defined by the theory of constituent structure. The second part of this formal characterization makes clear the idea that LD is (head-)adjunction under adjacency. Beyond these

formal properties of LD, another factor that potentially interacts with processes of this type is found in the idea that syntactic structures are interpreted at the interfaces cyclically. In section 4, I examine a specific way in which cyclic Spell-out interacts with PF operations and LD, based on an example from French prepositions and determiners. This case study paves the way for further investigation of cyclicity and the interface of syntax and PF.

1.1. Architectural Assumptions

The framework of *Distributed Morphology* is assumed here. This is a Non-Lexicalist theory of grammar, in which there is no generative Lexicon; this means that the derivation of complex objects takes place in the syntax, or in terms of operations that make reference to the output of the syntax (i.e., representations in the PF component).

I assume here a "dynamic" conception of PF, in which this component consists of an (ordered) set of computations that apply to the output of the syntax.[1] In the normal case, morphological structure is syntactic structure; i.e., PF operations that are "morphological" in nature apply to the structure that is the output of the syntax. This is part of a larger picture, in which structure generated by the syntax is modified in various ways at PF in accordance with that component's function of "packaging" syntactic representations for phonology.

The syntax derives hierarchical structures out of two types of terminal nodes. The first type, the *Roots*, correspond in many ways to the "lexical" or "open-class" vocabulary. The second type are functional heads (= "functional" or "abstract" morphemes). These feature bundles do not contain phonological representations in the syntax; rather, the phonological content of these nodes is added to them at PF. The process responsible for this is called *Vocabulary Insertion*. At PF, morphemes like T[past] (Tense with the feature [past]) have phonological material added to them, as specified in that language's *Vocabulary Items* (e.g., *T[past]* ↔ *-ed*). Following the application of Vocabulary Insertion, the node T[past] has a phonological matrix, e.g., T[past,*-ed*], where *-ed* is called the phonological *exponent* of this node.

[1] In some cases it appears that "PF" is used ambiguously in this sense and in the sense of "final phonological representation" (i.e., the output of the computations that are articulated in the dynamic view). Whatever the relevant notion of "final output" here might be, the point is that any derivational approach requires an account of what the different PF representations are and how they are derived.

A further set of operations in the PF part of the grammar linearize the hierarchical structure generated by the syntax. Local Dislocation, because it is defined in terms of linear adjacency, occupies an important position in the theory of linearization.

1.2. Linearization and Local Dislocation: Preliminaries

The starting point for the discussion of post-syntactic representations of linear order is the hierarchical representation generated in the syntax. An important assumption behind the project pursued in this article is that the structural notions relevant for PF representations and linearization are those defined by the constituent structure. For the purposes of this discussion, I take for granted certain notions that are operative in current research, such as "head," "complex head," and so on. The approach that is outlined here is exclusively structural. This means that, for example, the analysis makes no reference to the notion of "word" (prosodic or otherwise) as far as syntax and morphology are concerned—there are only structures ("head," "terminal," etc.) and their phonological interpretations. Whether or not there is ultimately the need for something like the *prosodic word* that cannot be defined structurally, it must be recognized that adding such an element to the ontology would constitute a significant modification.[2]

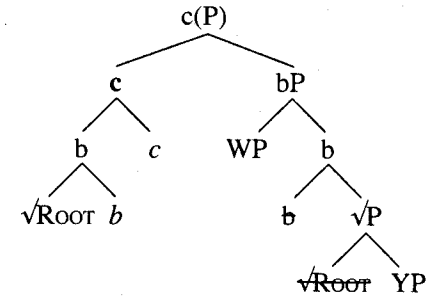
Some structures, and in particular the head, have an important status at PF. In particular, the normal case is for terminals within a single complex head to have an intimate phonological connection ("word-level phonology"). This aspect of the relationship between structure and phonology figures prominently in the discussion of section 3.

Some important structural notions are illustrated in (1). What (1) represents is a structure typical of head movement, in which a Root moves to functional head b, with the resulting complex then moving to functional head c:[3]

[2] For related questions concerning the general status of the prosodic hierarchy from the perspective of the kind of theory advanced here see Pak [28, 29].

[3] Unpronounced instances of moved elements are presented in strikethrough. I do not address the question of how it is determined which instance of a moved item is pronounced.

(1) *Hypothetical Structure*



One possibility discussed below is that the syntactically significant objects in this representation—the complex head(s) and the terminals within such heads—are themselves the objects that figure in statements of linear order. As a preliminary to this component of the discussion, the following definitions from Embick and Noyer [13] are relevant:

(2) *Definitions*

a. M-Word: (Potentially complex) head not dominated by further head-projection (cf. Chomsky [4] "H^{0max}")

b. Subword: Terminal node within an M-Word (i.e., either a Root or a feature bundle)

Illustrating with reference to (1), boldfaced **c** is an M-Word, while italicized *Root, b, c* are Subwords. Part of what is being investigated here is the idea that only these objects can be referred to by PF processes, such that other structure (e.g., the "intermediate" b in the complex head in (1)) has no status as far as the theory is concerned. This idea is clearly related to a parallel premise for phrase structure.

The structure in (1) shows a complex head created by one kind of syntactic affixation, head movement. In many cases, a complex head is created by a process that is not head-movement as typically understood. Instead, one piece is affixed to another under adjacency: Local Dislocation, in the terminology employed here. This operation is a descendant of "merger under adjacency" (Marantz [23, 24],and

A further set of operations in the PF part of the grammar linearize the hierarchical structure generated by the syntax. Local Dislocation, because it is defined in terms of linear adjacency, occupies an important position in the theory of linearization.

1.2. Linearization and Local Dislocation: Preliminaries

The starting point for the discussion of post-syntactic representations of linear order is the hierarchical representation generated in the syntax. An important assumption behind the project pursued in this article is that the structural notions relevant for PF representations and linearization are those defined by the constituent structure. For the purposes of this discussion, I take for granted certain notions that are operative in current research, such as "head," "complex head," and so on. The approach that is outlined here is exclusively structural. This means that, for example, the analysis makes no reference to the notion of "word" (prosodic or otherwise) as far as syntax and morphology are concerned—there are only structures ("head," "terminal," etc.) and their phonological interpretations. Whether or not there is ultimately the need for something like the *prosodic word* that cannot be defined structurally, it must be recognized that adding such an element to the ontology would constitute a significant modification.[2]

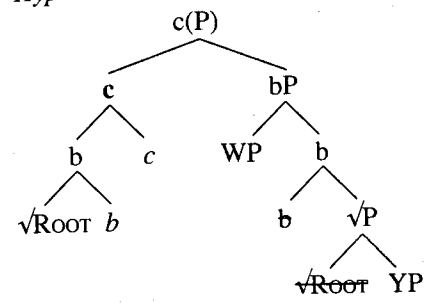
Some structures, and in particular the head, have an important status at PF. In particular, the normal case is for terminals within a single complex head to have an intimate phonological connection ("word-level phonology"). This aspect of the relationship between structure and phonology figures prominently in the discussion of section 3.

Some important structural notions are illustrated in (1). What (1) represents is a structure typical of head movement, in which a Root moves to functional head b, with the resulting complex then moving to functional head c:[3]

[2] For related questions concerning the general status of the prosodic hierarchy from the perspective of the kind of theory advanced here see Pak [28, 29].

[3] Unpronounced instances of moved elements are presented in strikethrough. I do not address the question of how it is determined which instance of a moved item is pronounced.

(1) *Hypothetical Structure*



One possibility discussed below is that the syntactically significant objects in this representation—the complex head(s) and the terminals within such heads—are themselves the objects that figure in statements of linear order. As a preliminary to this component of the discussion, the following definitions from Embick and Noyer [13] are relevant:

(2) *Definitions*

a. M-Word: (Potentially complex) head not dominated by further head-projection (cf. Chomsky [4] "H^{0max}")

b. Subword: Terminal node within an M-Word (i.e., either a Root or a feature bundle)

Illustrating with reference to (1), boldfaced **c** is an M-Word, while italicized *Root, b, c* are Subwords. Part of what is being investigated here is the idea that only these objects can be referred to by PF processes, such that other structure (e.g., the "intermediate" b in the complex head in (1)) has no status as far as the theory is concerned. This idea is clearly related to a parallel premise for phrase structure.

The structure in (1) shows a complex head created by one kind of syntactic affixation, head movement. In many cases, a complex head is created by a process that is not head-movement as typically understood. Instead, one piece is affixed to another under adjacency: Local Dislocation, in the terminology employed here. This operation is a descendant of "merger under adjacency" (Marantz [23, 24],and

related work) and ultimately of the "affix hopping" transformation of early generative grammar (Chomsky [3]).[4]

To take an example, the formation of English comparatives and superlatives shows distinct analytic and synthetic forms, depending on the phonological properties of the adjective involved:

(3) a. Mary is more intelligent than John.
b. Mary is smarter than John.

As discussed in Embick [14] (extending Embick and Noyer [13]), the affixation of the comparative morpheme (Deg) occurs under linear adjacency. As far as the syntax is concerned, both types of comparatives in (3) have an identical structure; this is represented in (4), where Deg is part of a DegP that is attached to *a*P:

(4) *Syntax of the Comparative*
$[_{aP}\ [_{DegP}$ Deg...] $[_{aP}$ Adjective...]]

At PF, there is a rule of Local Dislocation that affixes Deg to the adjective to produce a synthetic form when they are linearly adjacent, and when the adjective has the appropriate phonological properties. When the rule attaches Deg to the adjective as in (3b), Deg is realized as the "affix" *-er*, and the result is one "word" (i.e., a synthetic form). When the rule does not apply, Deg and the adjective are spelled out as two distinct "words," as in (3a).

In an abstract sense, Local Dislocation operations like that illustrated above take two separate elements—e.g., two M-Words—and create from them a single M-Word by affixing one to the other. Unlike the other types of syntactic affixation (head movement, Lowering), LD operates in terms of linear adjacency.

With reference to the structure in (1) and the definitions in (2), Embick and Noyer [13] define a "typed" conception of LD, in which the two categories of objects in (2) can only move with respect to one another:

(5) *Typing assumption on LD*: M-Words only dislocate with adjacent M-Words, and Subwords with Subwords.

[4] In addition to this, some approaches discuss the need for a "downwards" operation with the same locality properties as head movement (head to head of complement); this is Lowering in the terminology of Embick and Noyer [13].

This condition imposes restrictions on the application of LD, and is investigated in further detail below.

2. Some Properties of Local Dislocation

In the following subsections I examine two instances of LD, with reference to some particular properties that are important to the theory under discussion. The first example, involving the placement of the Latin enclitic *-que* 'and,' illustrates the typing assumption (5), along with some of its consequences. It also illustrates the importance of a *cyclic* conception of PF operations. The second case, drawn from Lithuanian, is based on the distribution of the "reflexive" morpheme *-si*. This case study shows that *-si*, a Subword with a requirement that something appear to its left, must satisfy this requirement within its own M-Word. This appears to be a case in which the "word" is special for linear relations, but illustrates and in fact follows from (5) as well.

2.1. Illustration 1: Latin *-que*

The Latin enclitic *-que* 'and' appears in simple conjunctions as an enclitic on the second conjunct, as the following examples show with elements of different grammatical categories:[5]

(6) a. Nouns:
diu noctu-**que**
'by day and by night'

b. Verbs:
vivimus vigemus-**que**
'we live and we flourish'

c. Adverbs:
bene pudice-**que** adservatur
'[She's] been chaperoned well and modestly'

[5] The properties of *que* are discussed in all standard reference works on Latin grammar; e.g. Kühner and Stegmann [21], Sommer [35], Ernout and Thomas [15], Leumann et al. [22]. Theoretical discussions of this clitic in terms that line up with the assumptions of this discussion are found in Marantz [24] and Embick and Noyer [13].

While the distribution of *-que* in these cases could be accounted for by a number of means, the broader generalization about the distribution of this element is that it attaches to the first head (in a sense to be made precise below) of the second conjunct, whatever that may happen to be. This pattern is seen clearly in cases in which the second conjunct is phrasal or clausal; "▼" marks the position between the two conjuncts, and "[" the beginning of the second conjunct:

(7) *Phrases/Clauses*

a. ...cum hac et praetoria cohorte cetratorum ▼
with this-ABL and official-ABL escort-ABL caetratus-GEN
[barbaris-**que** equitibus paucis
barbarian-ABL.PL-AND cavalry-ABL.PL few-ABL.PL
'with these and his official retinue of lightly-armed troops
and a few barbarian cavalrymen' C. *B.C.* II.75

b. ...▼[maius-**que** commodum ex otio meo
more-AND profit from idleness-ABL my-ABL
quam ex aliorum negotiis rei
than from others-GEN work-ABL.PL thing-DAT
publicae venturum.
public-DAT come-FUT.PART
'...and more profit will come to the republic from my
idleness than from the activities of others.' Sall. *J.* III.4

The distributional pattern illustrated above can be accounted for directly if attachment of *–que* occurs under linear adjacency. The first M-word of the second conjunct bears *-que*, no matter what that M-word is and no matter what syntactic configuration it may be in its own XP. A treatment in terms of syntactic movement would be problematic. The problem is that the process in question cannot be head-movement as standardly understood, since the elements hosting *-que* are not the syntactic heads of the conjuncts; nor can it be phrasal movement, since such an operation would not have the ability to move the "first word" of a conjunct and (evidently) attach it to a head.[6]

[6] Even if remnant movement were appealed to, there would be a question as to where the remnant containing the first word of the second conjunct was moving, and why this process had to affect a remnant containing one and exactly one word.

An adjacency-based treatment of *-que* is schematized in (8-9); X,Y,Z stand for M-Words, ⌢ for concatenation (see below), and ⊕ for affixation produced by LD:[7]

(8) *Structure (Partial)*
[CONJ [$_{Conj2}$ X...Y...Z]]

(9) *Local Dislocation*
CONJ⌢X → X⊕CONJ

A point that casts some additional light on this process is that many of the hosts of *-que* are internally complex, consisting of more than one piece. The verb *vigemus*, for instance, consists of a Root, a Theme Vowel, and an agreement node. However, *-que* is affixed to the entire M-Word ([$_M$ = M-Word boundary):

(10) a. Structure:
[que [vigemus]]
'...and we flourish'

b. Output:
vigemus-**que**

c. Internal structure:
[$_M$ que] ⌢[$_M$ [vig e] mus]

d. No Interpolation:
*vig-**que**-e-mus
*vig-e-**que**-mus

In order for Local Dislocation to generate the correct outcome, it is necessary to assume that when the M-Word *-que* is moved, it can only target another M-Word. This is important in the following sense: it is not the case that the M-Word *-que* attaches to the adjacent *piece*; this would predict the unattested interpolations seen in (10d). Rather, the object that is targeted by the moving M-Word *-que* is the entire adjacent M-Word. This is an illustration of the typing assumption: the M-Word is placed with respect to the following M-Word, not the following Subword.

[7] The syntactic status of the CONJ realized as *-que* depends on assumptions about the syntax of coordination, as well as assumptions about peripherality in LD (cf. Marantz [24] and section 3-4).

Some further aspects of the distribution of *-que* implicate a kind of cyclicity. With PPs headed by many monosyllabic prepositions, *-que* surfaces after the complement to P, not affixed to the preposition, as shown in (11a). This is not the case with all prepositions as (11b) illustrates (see Ernout/Thomas [15:120,454], Kühner/Stegmann [21, v.1:583, v.2:10]):

(11) a. in rebus-**que** 'and in things'
de provincia-**que** 'and from the province'
ad Caesarem-**que** 'and to Caesar'

b. circum-**que** ea loca 'and around those places'
contra-**que** legem 'and against the law'
sine scutis sine-**que** ferro 'without shields and without sword'

This pattern results from the operation of another process. In the (11a) cases, the relevant prepositions undergo an additional string-vacuous LD operation to become procliticized to the following noun:[8]

(12) $[_{\mathcal{M}}$ in$] \frown [_{\mathcal{M}}$ rebus$] \rightarrow [_{\mathcal{M}}$ in⊕rebus$]$

The process shown in (12) is an instance of a general process that affects the relevant prepositions. It occurs inside the PP and precedes the affixation operation that places *-que*. At the stage of the derivation when the rule affecting *-que* can apply, the derived M-Word *in⊕rebus* is adjacent to *-que*, so that LD affixes *-que* to this, and not to the preposition itself.[9]

In the way noted above, the analysis of *-que* illustrates the role of the Typing assumption on LD. In addition, the analysis of the

[8] Some additional evidence for this operation comes from Latin orthography, which occasionally treats monosyllabic prepositions as part of the same word as the complement; see Sommer [35:289], Leumann et al. [22:241], among others. Note that treating the preposition and its complement as a single unit does not imply that they form a single unit for all types of word-level phonology (e.g., syllabification as revealed in the calculation of heavy and light syllables in verse). The procliticized prepositions might simply behave phonologically like prefixes in compound verbs, although the situation varies somewhat across periods of Latin poetry, to judge from Sommer [35:289].

[9] There are certain cases noted in the works above in which the prepositions in question host *-que*, e.g., when there is contrastive stress or in certain fixed phrases. I take these to be cases in which the general rule behind (12) does not apply, for reasons that could be articulated in a comprehensive treatment of Latin prepositions.

P-cliticization facts implicates cyclicity. The interaction between P and N is determined before the dependency of *-que* is handled, something that follows from the idea that the dependencies inside the second conjunct are calculated prior to those involving *-que* itself.

2.2. Illustration 2: Lithuanian si

As discussed in Embick and Noyer [13], movement under adjacency is also apparently found within M-Words, resulting in what looks like "affix movement." This is found when a Subword surfaces in a linear position that is unexpected given its hierarchical location, or when a particular morpheme shows a second-position effect within an M-Word. For such cases, Embick and Noyer [13] hypothesize that Subwords undergo LD with respect to adjacent Subwords, in a typed fashion as per the discussion above. While there is some question about whether or not this operation can be treated as Local Dislocation in terms of a typed system like that discussed above (see section 1.2), there are some important observations to be made concerning "word-internal" interactions. In particular, it appears that satisfaction of a Subword's requirements is handled internal to the M-Word in which that Subword originates. Subwords do not seem to "escape" one M-Word and attach to another. The placement of the Lithuanian "reflexive marker" *–si* (Senn [34], Ambrazas [2], Nevis and Joseph [27]) is relevant to this question. While this is a complicated case, the basic point that I hope to establish here concerning the "no escape" property can be established directly.

In simple (unprefixed) verbs the reflexive morpheme (exponent *-si*) appears after the Verb-Tense complex (13a); in verbs with a single prefix, *-si* appears between the prefix and the verb (13b); and in verbs with two prefixes, *-si* appears between these two prefixes (13c):

13) a. laikaũ-**si** 'I get along'
b. iš-**si**-laikaũ 'I hold my stand'
c. su-**si**-pa-žinti 'to know [someone], to recognize'

One clear point about this distribution is that *-si* cannot occur word-initially; that is, it is suffix-like in the sense that it always has an element to its left. The complication, of course, is that *–si* does not always appear at the end of the verbal complex; rather, it shows a type of second-position behavior.

This distribution can be understood as follows. The morpheme *-si* is adjoined highest in the node containing the verb; in the case of finite verbs, it can be assumed that the verb undergoes head movement to T, and that the REFL morpheme is attached to this entire complex. The simplest assumption is that REFL is initially linearized as the leftmost element. Then, in accordance with the requirement that *-si* be suffixal, it undergoes LD with the adjacent Subword.[10]

For the purposes of this initial stage of the discussion, an important point is that *-si* is restricted to finding a host within the M-Word that it originates in. While elements may of course precede the verb in Lithuanian, it is never the case that *-si* simply suffixes onto one of these to satisfy its requirement, as in (14a); the action occurs within the M-Word containing *-si*, as in (14b):

(14) a. *àš-**si** ne-lenkiù
I-REFL NEG-bow.1S
'I do not bow.'

b. àš ne-**si**-lenkiù
I NEG-REFL-bow.1S
'I do not bow.'

The general point, of which this is one instance, is that it appears that a Subword cannot escape an M-Word in order to satisfy its requirements. That is, although *-si* requires something on its left, this requirement does not take into account elements outside of the M-Word to which *-si* belongs. The effect here is related in some sense to what is seen in some of the Latin examples, where moving *-que* does not interpolate inside of a complex element of a particular type. That is, the fact that M-Words do not move inside of M-Words and the fact that Subwords do not move out of M-Words, seem to be related to one another. As will be shown below, both the absence of interpolation into an M-Word and the absence of escape from inside of an M-Word can be made to follow from the same general condition on linear relations, in which it is basic statements of linear order that are typed.[11]

[10] For the distinct types of requirements at play here, see the different types of linearization statements employed in section 3.

[11] While Lithuanian *-si* occurs between prefixal elements and the verb, it does not interpolate between the verb and Tense/Agreement pieces, as discussed in Embick and Noyer [13]. The treatment of this effect within the present system relies on assumptions about the type system involved; see section 3.3.2. Other considerations—e.g., the possibility of having REFL begin as a suffix when there are no pre-verbal morphemes—might lead to possible treatments as well.

2.3. Synopsis

These cases of movement and others examined in the literature appear to be cases of affixation under adjacency. This leads directly to the question of what this means formally; i.e., in terms of the hierarchical and linear representations found at PF. A further point illustrated in the cases chosen above is that LD is subject to restrictions on its application. It does not, as far as I know, allow interpolation of an M-Word into an M-Word, or allow a Subword to escape an M-Word. These restrictions and the LD operation in the first place should be stated in terms of (and ideally follow from) the linearization representations that are required in the "normal" case.

3. Hierarchical Structure and Linear Order

The information involved in the linearization of syntactic structures is of different types, and relates to statements that are more and less abstract in terms of what they relate. One type of linear representation relates categories to one another. For example, statements of headedness of the type "X precedes its complement YP" relate members of the category X to a phrase, YP. Such statements encode generalizations that go beyond the properties of individual terminals. When X precedes YP this means that X appears to the left of the first element of YP, whatever YP may happen to contain. I assume that this kind of linear information is encoded in a process that for [X Y] generates either (X * Y) or (Y * X), where * is an operator for "is left-adjacent to." This information is generated for each (branching) node in the syntactic structure. While I make use of *-statements here as a means of stating headedness generalizations, the crucial part of the theory of Local Dislocation makes reference to another type of statement, involving concatenation. Whether other types of processes require *-statements is a question for further research; one possibility is that this step in the procedure can be eliminated.

Beyond the more abstract * relation, PF must impose further order on the syntactic terminals, since * does not specify a concatenation

of terminal nodes. By "terminal" at this stage I mean "M-Word"; I return to the ordering of Subwords below. Using $\frown$ for concatenation, this part of the linearization procedure produces statements like $(X \frown Y)$ and $(Y \frown Z)$.[12]

Finally, concatenated elements must be "chained" into a linear representation that can be employed by the input/output system; I will have little to say about this chaining step here (see Pak [29] for discussion centered on syntax/phonology interactions).

The steps described above are schematized in (15). I assume that both larger objects (M-words, phrases) and smaller objects (Subwords within M-Words) are subjected to the same procedures. For expository purposes, I continue to use $\oplus$ to indicate concatenation of Subwords as opposed to M-Words:[13]

(15) *Phrase Structure → Linear Order*

a. Syntactic Structure: $[_{XP}\, X\, [_{YP}\, Y\, Z]]$
Hierarchical representation. Relevant for e.g., *Lowering* (head lowers to head of complement; see Embick and Noyer [13] and section 4).

b. *; Adjacency: Represented as: $(_{XP}\, X * YP)$, $(_{YP} Y * Z)$
* = "is left adjacent to"; representation of headedness/adjacency of abstract objects (phrases, etc.)

[12] A linearization procedure with two-steps like * and $\frown$ is employed in Sproat (1985); cf. also Marantz (1984). A two-step procedure with different properties (relations) is sketched for some complex cases in Embick and Noyer (to appear)). It might be the case that certain phenomena require an analysis in terms of operators that have properties different from * and $\frown$.

[13] In some of the representations below I abbreviate with statements like those in (i):

(i) a. $*: (_{XP}\, X * (_{YP}\, Y * Z))$
b. $\frown: X \frown Y \frown Z$

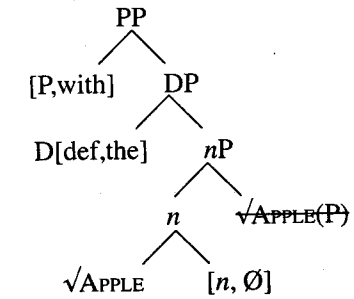
The statement of Local Dislocation operations involves only binary statements, for reasons that are discussed below. Ultimately, however, there is a question as to whether representations like those in (13) might be relevant for certain phenomena. In particular, it could be asked how * and $\frown/\oplus$ statements relate to phrasal phonology. One specific question concerns which types of linearization statements are referred by such prosodic rules. Pak [28, 29] follows earlier work in identifying different types of phrasal rules, and advances the further proposal that these rule types apply at different stages of the PF-derivation, i.e., operate in terms of different statements of linearization along the lines outlined above. Within such an approach it is an empirical question how much information appears in the formulation of the phonological rules in question (e.g., how big the objects in question are).

c. $\frown/\oplus$; Concatenation:
Represented as: $X \frown Y$, $Y \frown Z$; $a \oplus b$, $b \oplus c$
$\frown/\oplus$ = concatenation for M-Words/Subwords respectively

d. Chaining: Representation of all information in a linear sequence.

For illustration, consider (16) (phonological exponents of the functional heads are included for expository purposes):[14]

(16) *with the apple*



Focusing first on relations among larger objects (M-Words and phrases), the first stage of linearization introduces the *-operator, in a way that reflects general properties of English (e.g., PPs are head-initial, etc.):

[14] To this point, nothing has been said concerning the ordering of the linearization operations in (15) with respect to Vocabulary Insertion (VI). Embick and Noyer [13] hypothesize that Local Dislocation occurs "after" this process. The reason for this is that many LD operations are sensitive to phonological or morphological properties of specific Roots, which might not be present unless some VI has taken place (i.e., if one assumes late insertion of Roots). Importantly, the cases in question are cases in which it is the phonology of the target that is at issue. Whether or not the element moved by LD (the "affix") has been subjected to VI prior to LD is not always clear. It seems that there are cases in which an element moved by LD acquires its phonological form only after moving, but there are a number of factors to take into account in this. For instance, it might be that a concatenation statement like $X \frown Y$ makes Y close enough to X so that it would be visible for contextual allomorphy (e.g., indefinite article in English, if this involves two distinct Vocabulary Items).

In any case, some of the analyses below appear to support the idea that LD can move a node before VI occurs—see in particular the discussion of French in section 4—whereas e.g., Latin preposition cliticization appears to occur after P has undergone VI. The interaction between Local Dislocation and Vocabulary Insertion is worth examining in greater detail, although I will not do this here.

(17) *Stage 1*

LIN-*[n √(P)] → (n * √(P))
LIN-*[D[def,the] nP] → (D[def,the] * nP)
LIN-*[[P,with] DP] → ([P,with] * DP)

The representation that is the input to late stages of PF contains information about concatenated terminals. I assume that the procedure involved in determining such statements operates on representations derived by LIN-*, and has the following effects:

(18) For X(P) = [W_1 … W_n] and Y(P) = [K_1 … K_n], where W_i, K_i are M-Words, (X(P) * Y(P)) → (W_n ⌢ K_1)

Thus what it means for X to be next to Y when (X * Y) is that the final element of X is concatenated with the initial element of Y; see Marantz [23] for some discussion of the relevance of this procedure for clitic phenomena. With reference to (17), the second set of linearization statements are as follows:[15]

(19) *Stage 2*

LIN-⌢ [(n * √(P))] → (n ⌢ ~~√APPLE~~)
LIN-⌢ [([D,the] * nP)] → ([D,the] ⌢ n])
LIN-⌢ [([P,with] * DP)] → ([P,with] ⌢ [D,the])

One of the elements that figures in these statements, the n, is internally complex; i.e., it is an M-Word containing more than one Subword. Complete linearization requires linearization of the contents of M-Words as well. I assume that the procedures outlined above apply in the same way; from * statements, the system derives statements of concatenation between Subwords (represented with ⊕):[16]

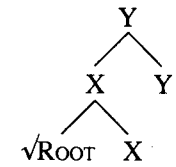
[15] There is a question about whether non-pronounced copies figure in the linearization statements of the type discussed here. It is possible that the system could be defined in either way, although the simplest view mechanically would be that unpronounced elements are present until eliminated by some further operation. This point connects with some other issues. There is the possibility that the locality conditions on contextual allomorphy might be statable in terms of concatenation, and that (at least some) morphemes with null exponents are "invisible" for this process; see Embick [10] for discussion.

[16] There is a question here about what range of structural possibilities are found within complex heads in the first place. While in many cases head-movement produces "simpler" structures compared to those found in the larger domain—i.e., structures that are derived by successive adjunction—there are cases in which more complex structures appear M-Word-internally as well (e.g. in the case of compounds like *butterfly net collector hat*).

(20) For X = [W_1 … W_n] and Y = [K_1 … K_n], where W_i, K_i are Subwords, (X * Y) → (W_n⊕K_1)

In (21-22) the linearization procedure is illustrated in a case where a Root combines with two heads X and Y:

(21) *Structure*



(22) *Linearization*

a. *: ((√ROOT * X) * Y)

b. ⊕: √ROOT⊕X, X⊕Y

With reference to the example above, this concatenation statement (√APPLE⊕n) is derived. There are two important assumptions that underlie the proposal developed to this point, each of which has implications for the restrictions on LD operations. The first is that the linearization procedure itself is typed; this is the source of the typing restriction on LD:

(23) *Typed Linearization Hypothesis*

Statements of concatenation are typed; i.e., they relate only elements of like type. There are at least two types: M-Words and Subwords. Where upper case X,Y are M-Words and lower case a,b are Subwords, linearization procedures generate two types of concatenation statements, X ⌢ Y and a⊕b. No such statements exist between objects that are not identical in type.

The second hypothesis is that the procedure derives (i) concatenation statements between M-Words; and (ii) concatenation statements between Subwords within a particular M-Word. It does not derive concatenation statements between Subwords of adjacent M-Words. Thus when an M-Word X containing Subwords a,b is linearized with respect to (right-)adjacent M-Word Y with Subwords c,d, the procedures above derive what is represented graphically in (24):

(24) *Graphically*

$[_M$ X] ⌢ $[_M$ Y]
| |
[a⊕b] [c⊕d]