Nat Lang Linguist Theory (2007) 25:1–37
DOI 10.1007/s11049-006-9002-9

ORIGINAL PAPER

Blocking effects and analytic/synthetic alternations

David Embick

Received: 25 April 2005 / Accepted: 20 February 2006 /
Published online: 11 January 2007
© Springer Science+Business Media B.V. 2006

Abstract A number of interactions in grammar are referred to as showing *blocking effects*, typically defined as cases in which the existence of one form prevents the existence of a form that is otherwise expected to occur. Patterns of analytic/synthetic alternation, in which two-word and one-word forms alternate with each other, have been taken to be instances of blocking in this sense. An example is found in the formation of English comparatives and superlatives, where, for example, the synthetic form *smarter* appears to block the analytic form **more smart*. Analytic forms are available in other cases (e.g. *more intelligent*), such that the interaction between the "one word" and "two word" forms is crucially at issue. This paper examines English comparative and superlative formation, concentrating on the question of how the morphophonology relates to syntax and semantics. A central point is that in the architecture of Distributed Morphology, these cases do not involve word/word or word/phrase competition-based blocking. Rather, blocking effects broadly construed are reduced to the effects of distinct mechanisms: (1) Vocabulary Insertion at a particular terminal node (morpheme), and (2) the operation of combinatory processes. The paper provides a detailed discussion of the latter type, showing that synthetic comparative/superlative forms are created post-syntactically by affixation under adjacency. Throughout the discussion, questions concerning the status of blocking effects in Distributed

An earlier version of this paper—in particular, the initial attempt to distinguish different types of so-called blocking effects and the outline of a treatment of comparative and superlative formation—was presented at Princeton University and the Coloquio de Morfosintaxis at the University of Buenos Aires, and parts were also discussed in my 2004 seminar at the University of Pennsylvania. I am grateful to these audiences for a number of helpful comments, which prompted a more detailed examination of comparatives along the lines presented here. For discussing this material with me at different points and commenting on earlier handouts or drafts I would also like to thank Rajesh Bhatt, Robin Clark, Morris Halle, Alec Marantz, Rolf Noyer, Marjorie Pak, and Maribel Romero. Finally, the article has improved because of the comments of NLLT reviewers.

D. Embick (✉)
Department of Linguistics, University of Pennsylvania, 619 Williams Hall
Philadelphia, PA 19104-6305, USA
e-mail: embick@ling.upenn.edu

Morphology, and those found in analytic/synthetic alternations in particular, play a central role.

1 Introduction

Approaches to grammar differ architecturally in terms of whether they posit a division of labor in the derivation of complex objects. For the purposes of this discussion, there are two major lines of research to consider. *Lexicalist* theories are theories in which (at least some) derivation of complex objects, and in particular at least some word formation, takes place in the *Lexicon*, defined for this discussion as a generative system distinct from the syntax. On the other side of this basic division, *Non-Lexicalist* theories are those in which all derivation of complex forms takes place in the syntax; there is no generative Lexicon.

In the treatment of blocking effects, the different analytical possibilities associated with or forced by the different theories are revealed clearly. The term 'blocking' is employed in Aronoff (1976) to refer to cases in which the non-existence of an expected form is supposed to result from the existence of another form. Since the initial application of this term, a number of patterns have been argued to be instances of blocking, including the patterns of comparative and superlative formation examined in detail in this paper. Such phenomena are of interest to grammatical theory because the treatment of blocking effects requires explicit assumptions about the combinatory mechanisms that create complex objects and the interaction of such mechanisms with other aspects of the grammar, such as the property of being on a list. Much of this paper is devoted to showing how these effects are implemented in the Non-Lexicalist theory of Distributed Morphology, which provides a perspective on these points that is distinct from that commonly found in the literature.

Aronoff's (1976) original formulation of blocking relies on the idea that words that are irregular in some respect are listed lexically, and that lexically listed words can block those derived by rule. With this intuitive association between blocking, wordhood, and the lexicon in the background, particular interest is generated by the phenomenon discussed at length in Poser (1992) in which words and phrases appear to interact in a way that implicates blocking. Employing examples from the English comparative and superlative, these are patterns of the following type: *smart, smart-er* versus *more smart*; *intelligent, more intelligent* versus *intelligent-er*; and so on. In cases like this, the factors conditioning the creation of such forms make it appear as if the existence of a particular 'one word' synthetic form precludes the existence of a particular 'two word' analytic form. At some level it must be asked whether there is blocking in cases of this type or not.

If the analysis of blocking envisioned by Aronoff is to be extended to these cases, then certain architectural modifications might be required. For example, Lexicalist theories posit an architectural difference between the creation of words and the creation of phrases. For this reason, such approaches appear to face problems in cases where analytic forms, which are supposed to be constructed by syntactic derivations, alternate with synthetic forms, which are supposed to be created by rules in Lexicon. If *smarter* blocks *more smart*, as Poser (1992) and others following him have proposed, then blocking must be extended out of the lexicon and into the syntax. Since interactions between these distinct components of the grammar are not permitted in (at least typical versions of) the Lexicalist architecture, the grammar must be set up so that

words can sometimes take precedence over phrases. The proposal advanced in Poser (1992) and subsequent work attempts to implement this proposal by allowing limited interaction between Lexicon and syntax—enough to allow words to block phrases under certain conditions. One could ask whether allowing even limited interaction between the 'word' and phrase systems is compatible with the Lexicalist program. However, my primary focus in this paper provides the groundwork for another question; namely, whether the blocking approach to such interactions makes the correct predictions about the range of environments in which word/phrase interactions occur.

While motivating some sort of change in perspective in Lexicalist frameworks, alternations between 'words' and larger syntactic objects (phrases) seem to be directly compatible with a syntactic approach to morphology; i.e., with the Non-Lexicalist view. In such approaches, a single system is responsible for the generation of all objects, whether they surface as 'one word' or 'two words'. The ultimate (morphophonological) packaging of a set of terminal nodes as 'one word' or 'two words' has the potential to shed light on the nature of the combinatory processes in the grammar and the conditions that govern their application, as is demonstrated below. However, while analytic/synthetic alternations provide a complication to the relationship between syntax and morphophonology, all of the relevant processes required for the analysis of such phenomena are found either in the syntactic derivation, or in PF, the component that interprets the output of narrow syntax. Another way of putting this would be to say that interactions of the analytic/synthetic type are not unexpected in the context of Distributed Morphology; however, expectations aside, the question is once again whether the architecture of this theory defines an analysis in which the range of analytic/synthetic interactions is properly circumscribed.

The salient difference between the approaches mentioned above is in terms of whether the interaction between *smarter* and *more smart* involves competition between words and phrases (blocking) in the first place. Theories that follow Poser (1992) say that it is blocking. The view that follows from Distributed Morphology is that this is not a case of blocking, and the bulk of this paper is devoted to showing exactly what this means. In addition to making this point precise, the analysis presented here lays the foundation for a comparison of different grammatical architectures. Implementing an analysis of word/phrase interactions as blocking along lines envisioned by Poser results in a theory with different properties from the type of approach advanced here. As I have stressed above, the most interesting question is ultimately whether syntactic theories like Distributed Morphology or theories with "words beating phrases" make the correct empirical predictions. Some comments clarifying what is at issue are presented below.

After outlining some architectural points later in this section, I illustrate how 'blocking' understood as an effect in which the existence of one (derived) form prevents the existence of another (derived) form does not exist in the approach of Distributed Morphology (Sect. 2). The bulk of the paper (Sect. 3) is devoted to a syntactic analysis of English comparative and superlative formation. This discussion moves beyond standard cases of comparative and superlative formation to address cases in which synthetic forms never appear, such as 'metalinguistic' comparatives like *John is more lazy than stupid*. It is shown that a treatment of the analytic/synthetic pattern must take account of the syntactic and semantic properties of comparatives. The discussion of Sects. 2–3 illustrates how the Distributed Morphology analysis of blocking effects provides a perspective on these phenomena that is quite distinct from what is found in most of the literature. Some implications of this view are discussed in Sect. 4.

Before examining blocking phenomena more closely, some general points are in order concerning the Non-Lexicalist architecture assumed here. The analysis that is developed below assumes the basic principles of Distributed Morphology, an approach that makes explicit claims about the structure of the Non-Lexicalist grammar (cf. Halle & Marantz, 1993 and subsequent work). In the default case, morphological structure is syntactic structure. To the extent that there are 'morphological operations', these are computations that take place as part of the PF component. According to this view, PF is an interpretive component that performs various operations on the output of the syntactic derivation; these operations eventuate in phonological/phonetic representations.

In the normal case, the PF operations that are relevant to typically morphological concerns are (1) Linearization, a set of processes that define linear order on the hierarchical structure generated by the syntax, and (2) Vocabulary Insertion, a process that provides phonological content to functional heads, which are assumed to be bundles of features without phonological content in the syntactic derivation. Linearization is discussed in detail later in the paper. Vocabulary Insertion is illustrated for the English past tense in (1). In this example, it is assumed that the English past tense involves a syntactic structure that contains the head T[past], whose phonological features are supplied at PF by entries ordered by specificity in the familiar way:

(1) Vocabulary Items for Past Tense (T[past])
T[past] ↔ -t/__ {$\sqrt{\text{LEAVE}}$, $\sqrt{\text{BEND}}$, ...}
T[past] ↔ -∅/__ {$\sqrt{\text{HIT}}$, $\sqrt{\text{QUIT}}$, ...}
T[past] ↔ -ed

The items in (1) are *Vocabulary Items* (VIs). When a VI wins a competition at a particular morpheme, it is at the expense of other VIs. An important point is that competition is restricted to single nodes (cf. Sect. 2). Notationally, I represent the output of the application of a VI to a node like T[past] with T[past, -ed], where the second component *-ed* indicates the phonological *exponent*.

The $\sqrt{\text{ROOTS}}$ (Roots) in (1) are elements of the open-class vocabulary. For present purposes I assume that Roots are present in the syntactic derivation (for discussion see Embick, 2000); that is, they are not subject to 'late insertion', as the functional heads like T[past] in (1) are.

Some additional PF operations perform minor readjustments to the syntactic structure under certain circumstances. For example, there are operations that affix (by head adjunction) one terminal to another under linear adjacency. Operations of this latter type are examined in detail below.

2 Blocking

Some basic observations and intuitions about blocking were outlined in my introductory remarks. A fundamental theoretical question concerns the scope of this phenomenon. Although a number of phenomena are referred to as instances of *blocking*, a closer examination reveals that the intuitive notion associated with this term might not pick out a single phenomenon as far as the grammar is concerned. In Aronoff's original formulation, blocking is "...the nonoccurrence of one form due to the simple existence of another" (1976:43). The context for Aronoff's proposal is an attempt

to explain that it is that the existence of *glory* that makes **glorious-ity* ungrammatical. The idea is that the fact that *glory* exists on a list in the grammar prevents the derivation of **glorious-ity*, in a way to be elaborated on below.

In other contexts, the term *blocking* has been invoked in explaining the putative deviance of e.g. *stealer* with respect to *thief* (for blocking in this sense see Giegerich, 2001, cf. Marantz, 2003 for a critical perspective). Finally, and more relevantly for the study of analytic/synthetic alternations, it has been claimed that e.g. *smart-er* blocks **more smart* (e.g. Poser, 1992; Kiparsky, 2005).[1]

In order to accommodate the distinct phenomena noted above under a single label, a generalized notion of blocking would take the form of something like (2):

(2) **Blocking (Abstract):** A case in which the existence of one form prevents the appearance of another form whose existence would otherwise be expected (all other things being equal).

It should be clear that (2) is not an analysis; instead it is a name for a particular kind of effect which may or may not have a uniform grammatical encoding. Importantly, (2) fails to specify precise definitions for the key terms—in particular the meanings of 'existence' and 'prevents'—that have to be articulated precisely in any satisfactory account. The empirical predictions of various theories are only revealed when the phenomena falling under (2) are given a formal analysis, as is made clear below.

One of the major goals of this paper is to articulate how blocking effects are accounted for in a syntactic approach to morphology. When different cases of blocking phenomena covered by (2) are examined in terms of the architecture for morphosyntactic derivations provided by Distributed Morphology, it is clear that there are distinct mechanisms at work. In this initial discussion I present preliminary considerations on effects of allomorphy (*tak-en* versus **tak-ed*), turning to a detailed study of analytic/synthetic effects later. I put aside consideration of "semantic" or "synonymy" blocking of the *thief*/*stealer* type (see Marantz, 2003; Embick & Marantz, 2006 for comments). For the two types of phenomena associated with blocking to be examined here, the mechanisms that are required are Vocabulary Insertion on the one hand and rules that specify movement operations on the other:[2]

(3) Two types of interaction

[1] Although Poser (1992) and Kiparsky (2005) do not do so, one could ask something similar about *more intelligent* versus **intelligenter*. At some level, *more intelligent* surfaces and **intelligenter* does not; it could be asked why it is blocking when one word wins out over two, but not vice versa. See Embick and Marantz (2006) for discussion.

The fact that blocking is not something that can be used to diagnose wordhood—i.e. it is not restricted to word/word interactions—is noted by di Sciullo and Williams (1987). After noting the English comparative/superlative and the Latin Perfect (see Fn. 2), it is noted that:

…blocking is not a phenomenon restricted to "words", though its exact applicability remains unclear. It may or may not have something to do with the notion of listedness. Despite current ignorance, we may at least conclude that the phenomenon itself cannot be used to support any thesis implicating listedness as an essential feature of the concept "word", because *whatever blocking is, it is not restricted to words and in fact operates across the word/phrase boundary*. (1987:14; emphasis mine)

This means in effect that blocking phenomena at the level of abstractness of (2) cannot be invoked in the definition of 'word', a necessary definition for Lexicalist theories. In the context of a syntactic approach like that advanced here, there is no need to single out the 'word' as a grammatically special object in this way.

[2] While they are distinct, (B1) and (B2) may interact in the following way: allomorphy of a particular head might depend on what type of complex head it appears in. That is, the application of a

Footnote 2 continued

(B1) *Rule application at a single node:* A more specific Vocabulary Item takes precedence over a less specific Vocabulary Item. For instance, in (1) insertion of the exponent *-t* at T[past] in the context of $\sqrt{\text{LEAVE}}$ prevents the insertion of the default exponent *-ed*, because of how Vocabulary Insertion works: the rule insertion of VI with the exponent *-t* takes precedence over the VI that inserts *-ed*. This is a general property of Vocabulary Insertion (see Halle, 1997 for a formulation).

(B2) *Combinatory application:* All complex forms involve syntactic structure. In some special cases, combinatory processes that create complex heads (or single *phonological words*) apply under a restricted set of circumstances. This is the case in the examples invoked above: *smart, smart-er, *more smart; intelligent, *intelligent-er, more intelligent*. This does not involve rule application to a single terminal node; rather, it involves a rule that is combinatory in nature, which means that at least two nodes (and thus larger structures) are implicated. Here the term *combinatory application* is employed, to stress the fact that this phenomenon crucially implicates the mechanics for manipulating pieces in a derived syntactic structure.[3]

Some further comments are in order concerning (B1) and (B2). First, as noted above, there is a sense in which (B1) involves 'competition' between the VIs associated with distinct exponents. This competition is restricted to the determination of the phonological form of a single node ("morpheme"). Continuing with the *leave/left* example, there exist on a list in the grammar two distinct VIs that insert *-t* and *-ed*. The VI that inserts *-t* when $\sqrt{\text{LEAVE}}$ is present prevents the insertion of *-ed*. At the same time, each of these VIs has an independent existence in the grammar, and each plays a role in deriving the form of some grammatical objects.

What this means is that competition is restricted to one aspect of the derivation of complex forms, namely the consideration of the phonological form of single nodes (morphemes). There is no competition at the level of outputs, so that larger objects like "words" do not compete with one another. Consider, for example, *tak-en* versus **tak-ed*. *Tak-en* exists only as the result of a particular derivation, and has no independent existence on a list like the VI that inserts *-(e)n* in certain participial structures does. **Tak-ed* is not derived at all; i.e. it could only exist if the grammar were altered.

combinatory process of type (B2) could place a head in a structural position that meets the structural environment for a Vocabulary Item whose application would not be triggered if (B2) did not apply. So, for instance, in Latin, the Perfect tenses involve a head Asp[perf]. This head's phonological forms depend on whether or not it appears in a 'verb' (synthetic form) or a 'participle' (analytic form): *laudā-v-ī* versus *laudā-t-us sum*. See Embick (2000) for details.

[3] It should be noted that application or non-application of a rule could be relevant for things that are not pieces, such as the *Readjustment Rules* that perform phonological changes in particular morpho-syntactic environments. Rules of this type are responsible for "stem-changing", like the process that creates *broke-∅* in the past tense environment for $\sqrt{\text{BREAK}}$. The head T[past] has a zero exponent in the context of this Root and some others. The morphophonological change is the result of a Readjustment Rule. The principle here is the same: these rules apply when their structural description is met. Some conceivable examples (e.g. **Yesterday Mary break-∅ a record* or **Every week Mary broke-s a new record*) involve mis- or non-application of the required rule, and are thus not generated by the grammar. That is, the first example involves non-application of a required Readjustment Rule, the second over- or misapplication. In cases like this there is a problem with rule application, even though the rule is not a combinatory rule in the first place.

For a perspective on such 'stem-changing' processes in the context of a piece-based approach to word formation like Distributed Morphology see Embick and Halle (2005).

In the case of (B2) there is no competition between rules that exist on a list; nor are (output) options compared. Instead, there is a combinatory rule that applies when its structural conditions are met. There is therefore no sense in which *intelligenter* and *more intelligent* exist on lists that are consulted for insertion, nor does the grammar generate both *intelligenter* and *more intelligent* and select a winner. Rather, the syntax and PF generate a structure which, after Vocabulary Insertion, receives the phonological form of *more intelligent*.

For purposes of illustration it is instructive to consider in this context the sentences in (4):

(4)a. Where did John go yesterday?
b. *Where John did go yesterday?
c. *Where John went yesterday?

The examples (4b,c) are by hypothesis cases in which T-to-C movement has not occurred. These are not generated by the grammar. It is not the case that the existence of (4a) *blocks* the existence of the examples in (4b,c). That is, (4a) does not exist on a list, such that it competes with either of (4b,c) and wins, thereby blocking (4b,c). Rather, the grammar of English does not generate (4b,c) in the first place. The relationship between analytic and synthetic comparatives must be understood in these terms. Within Distributed Morphology, cases falling under (B2)—"Poser Blocking" in the terminology of Hankamer and Mikkelsen (2005)— are not instances of blocking: they involve a single structure in which a rule sometimes creates a synthetic form (complex head). This is made precise in Sect. 3.

There remain to be addressed some broad concerns associated with the project advanced here. A general question is whether an approach that implements (2) in terms of *distinct* processes like those identified in the discussion of (B1-2) involves a loss of generalization. Given the vagueness of (2), it should be clear that the reduction to distinct mechanisms is actually more precise than the apparently 'uniform' statement that is being replaced. This point is clear when we consider how the finer partitioning of blocking phenomena has parallels in other domains. A case in point concerns the manner in which movement phenomena are handled by the grammar. While many phenomena fall under the broad heading of 'movement', analyses of derivations involving movement in the general sense posit distinct mechanisms for subcases: XP-movement, head movement, post-syntactic movement, and so on (not to mention phonological cases such as metathesis). While all of these are 'movement' in an abstract sense—at a very broad level of description like that found in (2)—they are, under further analysis, cases that involve distinct grammatical mechanisms. There is no loss in generalization in replacing a notion of 'movement' applicable to all of these cases with different mechanical processes, if this is what the phenomena and their theoretical treatment warrant. The same is true in the domain of 'blocking phenomena', except that the reanalysis being entertained here is a bit more extreme—blocking in the sense of "one word/phrase exists and therefore another word/phrase does not" is eliminated, not partitioned. Ultimately this is something that is going to be decided empirically: either the theory that implements distinct processes makes the correct predictions (when compared to a "uniform" approach), or it does not.

The analysis advanced here involves a change in perspective in the way that blocking effects are understood. Many analyses of such effects, including the one that Aronoff (1976) presents, rely on the idea that it is the *listedness* of some form that is

the operative factor in determining a case of blocking. In competition for expression of some meaning, the listed forms win out over the non-listed competititors. Making this precise, Aronoff's (1976) implementation of the blocking effect involves the idea that words have associated with them particular slots, such that the presence of a form in a slot can preclude the derivation of another form that would fill the slot in question. Speaking of *glorious*/*glory*/**gloriosity*:

> We can account for the distribution ... simply by appealing to the fact that *+ity* derivatives of *Xous* adjectives must be listed in the lexicon. We may assume that the lexicon is arranged according to stems, and that for each stem there is a slot for each canonical meaning, where "canonical" means derived by regular rules...Let us furthermore assume that for each stem there cannot be more than one item in each meaning slot. If the *+ity* nominals are entered in the lexicon, then when we make up such a form we put it into the slot for abstract nominal for its stem. However, when there is already a nominal in the stem in question, then there is no room for the *+ity* nominal; it is blocked by the already occurring nominal. (1976:45)

This type of analysis defines a conception of the blocking effect in which the relevant objects that are competing are words. The interaction responsible for the occurrence and non-occurrence of certain forms is at the level of forms that exist on lists, more precisely, in paradigmatic 'cells' defined by their meaning. In order to account for analytic/synthetic alternations in these terms, a blocking-based account would have to allow words to block phrases under certain circumstances (Poser, 1992; Hankamer & Mikkelsen 2005; see Sect. 3.5 for comments).

The discussion of (B1-2) above implements a theory of blocking effects in which the explanatory burden is divided: it is (B1) at the level of the morpheme, and (B2) placed on the generative processes. If the analysis of such phenomena in terms of the mechanical processes identified above is correct and can be generalized, then the clear implication for blocking as conceived in this way is that there is no blocking as defined above in Distributed Morphology (see Embick & Marantz, 2006 for a more general discussion along these lines). In cases like (B1) all of the competition is restricted to which of a set of existing Vocabulary Items wins at a single node. In cases like (B2), what is at issue is whether or not a combinatory process applies. But there is no sense in which two forms are competing with one another for existence; complex forms exist only as the outputs of derivations, and not on lists that are accessed in the course of derivations.[4]

Some consequences of this view are taken up in Sect. 4 below. A first step in advancing our understanding of these issues is to implement in detail the syntactic approach to analytic/synthetic alternations, and the next section is devoted to this.

[4] Terminologically, some work uses the term *blocking* for the interaction of rules at a specific position—e.g., 'the rule inserting *-t* blocks the rule inserting *-ed*', but in order to avoid confusion I will eschew this use.

3 English comparative/superlative formation

At various points above I have referred to the formation of comparative and superlative forms in connection with what has been called 'combinatory application' (B2) above. This section develops an explicit account of comparative and superlative (henceforth C/S-) formation, in which (1) both analytic and synthetic comparatives have a uniform syntactic structure; and (2) the process responsible for creating 'affixed' comparative or superlative forms like *smart-er* or *smart-est* operates under string adjacency. The overarching theme is that this and related processes can be understood in terms of syntactic locality conditions. After examining the essence of an adjacency-based treatment in Sect. 3.1, metalinguistic comparatives like *John is more lazy than stupid* are treated in detail in Sect. 3.2. Such cases are important because they demonstrate that surface linear adjacency is a necessary but not sufficient condition for the creation of a synthetic form. Supplementing this, Sect. 3.3 provides an outline of a mechanical system required for the linearization of syntactic structures, something that is necessary for the formulation of an adjacency-based view. In Sect. 3.4 cases in which an adjective is pre-modified by an adverb, such as *John is amazingly smart*, are employed as a window on the structural locality conditions that affect C/S-formation. If affixation of the Deg element occurs under string adjacency, then it should not be possible if the adjective is pre-modified; Sect. 3.4 shows that this prediction is correct. A synopsis of this analysis and some general comments on blocking are provided in Sect. 3.5.

3.1 Background and preliminaries

The formation of synthetic forms like *smarter* as opposed to analytic forms like *more intelligent* is conditioned by a number of factors. An initial point about C/S-formation is that there is a prosodic condition on the host of *-er/-est*, such that synthetic forms are possible only with 'short' adjectives:[5]

(5)a. John is smart-er than Bill.
b. John is mo-re intelligent than Bill.
c. *John is intelligent-er than Bill.
d. #John is mo-re smart than Bill.

The prosodic condition results in synthetic forms for monosyllabic adjectives, and analytic forms for trisyllabic ones. The situation with disyllabic adjectives is complex, with both types attested. Since the focus of this paper is not on the metrical conditions involved here, I will simply refer to the adjectives with synthetic forms as "short in the relevant sense".

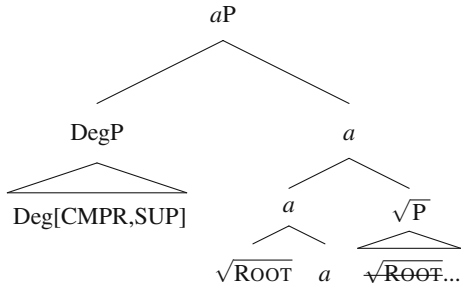
[5] I will discuss the status of *more smart* below; for the moment, it is simply treated as deviant.
The phonological conditioning factor on this process, in comparison with other cases of movement in which phonology does not play a role, has been amply noted in the literature, particularly in connection with other effects ('bracketing paradoxes'; see Pesetsky, 1979, 1985; Sproat, 1985; Marantz, 1988 and much subsequent work).

In treating the prosodic factor as something that is visible for the process creating synthetic forms, I am putting aside some further factors. For instance, for certain adjectives it appears to be the case that there is variation among and within speakers as to whether or not the comparative or superlative is analytic or synthetic. In addition, there might be additional reasons why adjectives that fit the prosodic condition in principle are not synthetic for other reasons (e.g. some speakers seem to generate *more clear* rather than *clearer*).

A second point is that although they are obviously different in form, the two types of comparative or superlative like (5a,b) receive the same interpretation. Therefore the starting point for a syntactic analysis of comparative/superlative formation is the assumption that there is a single syntactic structure underlying all comparatives and superlatives of the relevant type. This assumption allows for the most direct statement of the relationship between syntax and semantics in C/S-formation, which is uniform.[6]

Concerning the structure, I assume with Bhatt and Pancheva (2004) that DegP is attached to the phrase headed by the adjective (*a*P), an assumption that allows them to account for a substantial range of facts concerning the relationship of the Deg head to the adjective and to *than*-clauses, and, in addition, facts concerning certain aspects of the scopal behavior of Deg; see Bresnan (1973), Kennedy (1999), Lechner (1999,2004), Heim (2000) and much related work for relevant perspectives on the syntactio-semantic questions raised by comparatives. The structure-adapted slightly to conform with additional assumptions of this paper—is presented in (6):[7]

(6) Structure of Comparative/Superlative



DegP is headed by Deg[CMPR] (comparatives) or Deg[SUP] (superlatives). With reference to (6) and related structures, the term 'adjective' is used as an abbreviation for a Root combined with a category-defining functional head *a*. Notationally, I will not include reference to category-defining heads like *a* in some of the structures below, for the sake of simplicity. Instead, as a way of abbreviating this, I employ e.g. $[_A$ SMART$]$ for $[_a$ $\sqrt{\text{SMART}}$ $a]$ (similarly with "AP" rather than *a*P). When explicit details are required, I will represent the entire structure.

For a syntactic approach, the most salient question in light of the discussion of the previous section is what mechanical process creates synthetic C/S forms. The fact that

[6] A uniform syntactic treatment of Deg underlies most of the syntactico-semantic treatments of comparatives that I am aware of, where much of the emphasis is on constituency questions and on the interpretive properties of Deg, in particular scopal effects. If, on the other hand, analytic forms are created syntactically while synthetic forms are lexical, it is hard to see how the scopal behavior could be stated uniformly. If 'words' are opaque for syntactic operations, then even QR could not produce the desired effects, because it could not affect Deg that is attached lexically. Naturally it would be possible to achieve the desired effects through distinct mechanisms, but the question then would be why it should be necessary to make this move. This point has not, to my knowledge, been addressed in any Lexicalist treatment. For some discussions of the scope issues, see the references in the text and the references cited there. Some comments on lexical attachment of Deg are advanced at the end of this subsection as well.

[7] Here the $\sqrt{\text{ROOT}}$ moves to the category-defining head (*a*), as is generally the case with Roots. Bhatt and Pancheva take the DegP to be the specifier of the adjectival phrase, for them AP. I have represented it in a similar fashion in (6). No reference is made to the position of the subject of predication.

the process creating synthetic forms shows direct sensitivity to certain phonological properties of the host makes the question a difficult one, since for the most part movement operations as approached syntactically are not assumed to be conditioned by such factors.[8]

The implications of the phonological conditioning of C/S-formation is touched on in Embick and Noyer (2001), where the movement process is taken as applying late in PF after relations of linear adjacency have been imposed on the hierarchical structure derived by the syntax. This position has a theory-internal motivation, under the assumption that phonological properties cannot drive movement operations (or be referred to by movement operations) until this stage of the PF derivation. Concentrating on the mechanics of that proposal, rather than on correlative attempts at a motivation, the type of operation that converts the Deg head and the Adjective into a synthetic form is referred to as a *Local Dislocation*. A Local Dislocation applies under linear adjacency to 'affix' one element to another, as shown schematically in (7); this operation is (or is a form of) *merger under adjacency* (Marantz, 1988 and related work). When an element *X* undergoes Local Dislocation with *Y*, it is attached inside the complex head *Y*. The '*' is used for adjacency between heads and phrases, while the operator '⊕' is used to represent the relationship of adjacency that obtains within complex heads:[9]

(7) $(X * Y) \longrightarrow (Y \oplus X)$, or $(X \oplus Y)$

This initial formulation from Embick and Noyer (2001) leaves unclear some facets of linearization that are worked out below. At the level of abstraction assumed for (7), a Local Dislocation affects an element X and targets a Y of the same type as X, with X being affixed to Y. More precisely, this treatment assumes that a difference obtains between complex heads as a whole (*M-Words*), and the terminals that appear within these (*Subwords*). The M-Word is defined as follows:

(8) A node *X* is a *Morphosyntactic Word* (M-Word) iff *X* is the highest zero-level projection of X (cf. Chomsky (1995:245) "$H^{0_{max}}$").

The M-Words are simple heads (terminals) or 'complex heads', and the Subwords can be defined in these terms:

(9) *X* is a *Subword* iff *X* is a terminal node in an M-Word (i.e., a bundle of features)

The rule that creates synthetic comparatives and superlatives affixes Deg to the adjective. It operates on two M-words to create a single M-word, in the way specified in (10) for the synthetic comparative form *smart-er*; in this representation, the bracketing has been eliminated for ease of exposition:

(10) $\text{Deg[CMPR]} * [_A \text{ SMART}] \longrightarrow [_A \text{ SMART}] \oplus \text{Deg[CMPR]}$

[8] In fact, the idea that syntax cannot see any specifically phonologial properties has been proposed in different theoretical frameworks, and might have some motivation. The invisibility of phonology for syntactic operations can be stated architecturally in certain versions of Distributed Morphology (although it need not be); see Marantz (1995), Embick (2000), and Embick and Halle (forthcoming) for pertinent discussion.

[9] String-vacuous Local Dislocation is employed in the analysis of certain types of cliticization, and in the creation of morphophonological domains.

Given the phonological character of the constraints on this process, it must be assumed that the adjective is subjected to phonological rules that define its metrical structure prior to the operation that affixes Deg[CMPR].[10]

Continuing at a higher level of abstraction, the general rule of which (10) is a specific instance is formulated provisionally in (11), a language-specific rule of English:[11]

(11) English C/S Local Dislocation (Provisional)
Deg[CMPR,SUP] * $[...X...]_A \longrightarrow [...X...]_A \oplus$Deg[CMPR,SUP]
where the phonological form of $[...X...]_A$ meets the prosodic condition

The rule (11) is an initial formulation of the rule responsible for creating synthetic C/S-forms, one that summarizes some aspects of the sketch in Embick and Noyer (2001). Articulating different aspects of (11) and its relationship to syntactic structure occupies the greater part of this section, and leads to many points of interest.

As it stands, the rule (11) applies when two different types of conditions are met:

(12) **Condition 1 (C1):** The prosodic condition: the process applies only to 'short' adjectives, perhaps with certain exceptions as well.
Condition 2 (C2): Deg[CMPR/SUP] must be linearly adjacent to the adjective (see below).

To these can be added a third:

(13) **Condition 3 (C3):** The structure must be correct, in a way to be specified immediately below.

Condition 3 expresses the observation that linear adjacency between CMPR or SUP and an adjective is a necessary but not sufficient condition on the creation of a synthetic form out of these two pieces. As noted by di Sciullo and Williams (1987), among others, there are cases of surface linear adjacency in which synthetic forms are impossible:

(14)a. It is more hot than humid.
b. *It is hott-er than humid.

As is discussed below, the syntax of such 'metalinguistic' comparatives differs from that found with normal comparatives. The absence of synthetic comparatives follows from this structural difference in such a way that Conditions 2 and 3 are effectively non-distinct, in that in the metalinguistic comparative Deg and a potential host are never adjacent in the relevant sense.

It should be noted that there are other factors that come into play in determining the acceptability of comparatives and superlatives beyond (C1-3); the operative

[10] This is compatible with both "early insertion" and "late insertion" treatments of Roots. In the former case, the Root is present from the beginning, and it can be assumed to have undergone phonological processes after combining with *a*. In the latter case, it is standardly assumed that insertion proceeds from the Root outwards. Thus the Root is inserted and phonologized prior to consideration of various possibilities for Deg. In line with this, I assume here and in the detailed analysis later that Local Dislocation applies after Vocabulary Insertion has applied to nodes structurally inside of Deg, i.e. *a* and perhaps the Root. This is along the lines suggested by Embick and Noyer (2001). There is a further question about whether VI has occurred at Deg prior to Local Dislocation; for the points addressed in this paper, things can be made to work either way.

[11] This formulation does not address the question of whether the rule should be extended to cover adverbials.

word here is *acceptability* and not grammaticality, since a distinction must be made between syntactic and semantic well-formedness. For instance, the Root surfacing as an adjective must be capable of forming comparatives or superlatives in the first place by virtue of having the correct semantic properties. This is not a condition on the application of the process that creates synthetic forms in the same way that (C1-3) are meant to be. Rather, it expresses the claim that forms like *deader* are deviant (to the extent that they are actually deviant; consider *John's party was even deader than Bill's*) for reasons that have nothing to do with the creation of synthetic forms: there is a semantic problem with the combination of the Root and the structure that it appears in, and this has nothing to do with packaging as one or two words.[12]

Turning to some initial derivational details, the account sketched above holds that when the conditions in (C1-3) are met, Local Dislocation as defined in (11) applies and affixes Deg to the adjective. When the conditions are not met, a support operation inserts *mo-* to support the Deg head.[13] The two different types of derivation are illustrated in (15) and (16). The right arrow in the derivations below is used as a cover for different operations that are indicated to the left of the colon on each line (recall that the output of Vocabulary Insertion is given as [X, -x], with -x the phonological exponent of the node with X). For expository purposes, I have presented the output of Linearization as (Deg[CMPR] * ADJ), which simplifies the bracketing. Recall in addition that the Local Dislocation rule makes reference to the properties of the Root or Root/*a*, which have been operated on phonologically such that metrical properties are visible to the Local Dislocation rule (see fn. 10):

(15) Synthetic Form
  a. Syntax: [$[_{\text{DegP}}$ Deg[CMPR]] [$[_{A}$ SMART ...
  b. Linearization: (Deg[CMPR] * ($_A$ SMART))
  c. Local Dislocation:
     (Deg[CMPR] * ($_A$ SMART)) .... $\longrightarrow$ (($_A$ SMART)$\oplus$Deg[CMPR])...
  d. VI at Deg[CMPR]:
     (($_A$ SMART)$\oplus$Deg[CMPR]) $\longrightarrow$ (($_A$SMART)$\oplus$Deg[CMPR, -er])

(16) Analytic Form
  a. Syntax: [$[_{\text{DegP}}$ Deg[CMPR]] [$[_{A}$ INTELLIGENT ...
  b. Linearization:
     (Deg[CMPR] * ($_A$ INTELLIGENT))

[12] A reviewer suggests *primer* as a comparative adjective with little hope of salvation by coercion. Note that if any of the relevant properties of adjectives have a structural reflex—i.e., if the differences among some classes of adjectives involves a difference in structure or the features found in the structure—then this would amount to a condition of the type described by (C3). I am not aware of any arguments for this in the present case, however.

[13] Some additional details about the support process might be examined further. For instance, there is a question in (16c) about how to treat the *mo-* that is inserted here; one possibility is that first a (non-phonological) piece is attached to Deg[CMPR], and that it is this piece that has the exponent *mo-* added to it. It would also be possible to treat *more* with Vocabulary Insertion alone, so that *more* and *-er* would be two different allomorphs of Deg[CMPR]:

(i) Deg[CMPR] $\leftrightarrow$ more

Something in the VI inserting either *more* or *-er* would have to specify the structural environment (Subword status for *-er*, for example). I see no basis for choosing between these alternatives at present.

c. VI at Deg[CMPR] (recall fn. 13):
(Deg[CMPR] * ($_A$ INTELLIGENT))... ⟶
(Deg[CMPR,(mo)-er] * (($_A$ INTELLIGENT)...

The relevant processes are specified in further detail below, but these derivations suffice to illustrate what it means to have a uniform syntactic analysis of analytic and synthetic C/S-forms, in the terms sketched in Embick and Noyer (2001). In all instances, Deg has a uniform syntactically defined distribution, which is independent of the processes that apply at PF to yield analytic or synthetic forms.

Unlike head-movement or *lowering*—the process that combines T and *v* in English—Local Dislocation operates in terms of linear adjacency. Therefore the treatment that I have just outlined predicts that a synthetic form cannot be derived when anything intervenes linearly between the position of Deg and the adjective. In Embick and Noyer (2001) it is suggested that this kind of intervention is found in cases in which an adverb pre-modifies the adjective:

(17) John and Bill are both [amazingly smart].

In such cases, the synthetic form does not occur:

(18)a. Mary is the mo-st amazingly smart person.
b. *Mary is the amazingly smart-est person.

The examples in (18) show that when SUP scopes over the adverb and adjective, the analytic form (18a) must appear. In this type of example, the reverse scope of SUP and the adverb is not possible; this is shown in (19), where *intelligent* is used to avoid confounding factors:

(19) *Mary is the amazingly most intelligent person.

In examples with adverbs and adjectives, it seems that the formation of synthetic superlative forms is not possible when there is an adverb appearing linearly between Deg and the adjective. The same point can be established with the comparative, it is just slightly more complicated because the analytic version is grammatical, but with a different scopal reading:

(20) John is amazingly smarter than Bill.
=The degree to which John is smarter than Bill is amazing

That is, the scope here is [amazingly [Deg smart]], not [Deg [amazingly smart]] as it is in (18). In some cases it is more difficult to interpret an adverb in this higher position, e.g.:

(21) John is dangerously crazier than Bill.

This example is difficult to interpret at a first pass, because the reading derived from [dangerously [Deg crazy]] does not lend itself to as natural an interpretation as [amazingly [Deg crazy]] does. To the extent that it is possible to interpret the extent to which John is crazier than Bill as dangerous, (21) is acceptable.

In cases like *Mary is more [amazingly smart] than Bill*, the element that Deg is in a semantic relationship with is phrasal, the AP *amazingly smart* which contains an adverbial. For the purposes of Local Dislocation, the adverb follows the Deg head, so that the creation of an *-er* suffixed form of the adjective is not possible; based on the discussion above, what this means is that the representation in (22) does not meet the environment for the rule (11):

(22) (Deg * (ADV * ADJ))

As a result of the Local Dislocation rule not applying, the structure is subjected to the rules that eventuate in an analytic comparative. If correct, this line of reasoning would provide support for the structure and adjacency-based analysis developed here. For this reason cases of this type are examined in greater detail in Sect. 3.3.

The outline above presents some of the essential points of an analysis in which Deg is uniformly syntactic and subject to Local Dislocation. Two additional points specify some aspects of this analysis in further detail. First, concerning the linear aspect of the process that creates synthetic forms, a comparison between the Local Dislocation account can be made with one based on *head movement*, the syntactic process that creates complex heads. In the absence of arguments to the contrary, head movement must be considered as the operation responsible for synthetic forms, as it is required in other domains, and, all other things being equal, movement at PF of the Local Dislocation type involves syntax/morphology mismatches that should be avoided if possible. In the case at hand, there is evidence that the head movement analysis does not work. The process of 'head movement' in the domain of e.g. verb movement is conceived of as a process that skips adjoined or specifier material. For C/S-formation, head movement of this type would derive synthetic forms in which the adjective has moved higher than the adverbial, e.g. **Mary is the smart-est amazingly t person in the class*; this is clearly unacceptable. Thus even when it is assumed that Deg is in the correct position for head movement—i.e, even if Deg were a head taking an AP as its complement—it does not seem to be the case that there is evidence for a head movement analysis. Even if one were to stipulate that adverbs blocked head movement in this case, something would of course have to be said about the many cases in which adverbs appear to be invisible for this process.[14]

A second point is that there is no general prohibition against the derivation of synthetic forms in cases in which Deg is associated with a phrase (and not a 'simplex' adjective). That is, it might be thought that synthetic forms are created only when Deg applies to the adjective itself, and not some larger object like [ADV ADJ]. The idea behind this view would be that Deg can only attach to 'words', not to phrases. If this view were correct, it would account for the formation of analytic C/S forms in the *amazingly smart* cases in what could roughly be called semantic terms: synthetic forms can only arise when Deg gets together with a single 'word' that it is associated with semantically.[15] However, there are cases showing that this position is incorrect; this is seen in the fact that adjectives that take complements are capable of forming synthetic comparatives, even though it is the entire phrase that Deg applies to

(23)a. John is prouder of his son than Bill is.
[Deg [proud of his son]]

b. John is likelier to win the race than Bill is.
[Deg [likely PRO to win the race]]

c. John is quicker to point out problems than Bill is.
[Deg [quick PRO to point out problems]]

[14] If "Lowering" of Deg to the Adjective were possible, then *amazingly smartest* should be possible with Deg scoping over [ADV A]. This is not the case. It seems that the adverb is truly visible for the process that affixes Deg to the adjective (compare e.g. Bobaljik (1994), where it is suggested that adverbs are invisible for adjacency-based movements of a particular type).

[15] Note by way of contrast that on the approach adopted here it is always the case that Deg attaches to something phrasal.

The treatment based on structure and adjacency can account for the presence of the synthetic form here in the same way that it does for e.g. *smart-er*. Both the cases with adverbials and the cases in (23) are significant for comparing the syntax/PF approach here with alternatives, as noted in Sect. 3.5 below.

Elaborating on the point that Deg scopes over phrases, an anonymous reviewer suggests that a Lexicalist counter to the point about *prouder of his son* would be to hold that CMPR suffixes to *proud* directly in (23a) because it does not need to scope over the entire phrase *proud of his son*. The general principle that motivates this account is that CMPR or Deg elements in general are affixed to whatever they must scope over. Further consideration of this alternative seems to favor the syntactic approach over Lexical affixation once again.

This point arises in a consideration of ellipsis and its resolution (although cf. Fn. 16 below). Consider (24):

(24) John is prouder of Bill than Mary is Δ.

Here the ellipsis is resolved to [proud of Bill] (for simplicity, I am not including the degree variable). According to the syntactic account, the structure of (24) is given in (25); this is the syntactic structure prior to the PF operation that affixes Deg:

(25) John is CMPR [proud of Bill] than Mary is Δ.

There is strict identity between the antecedent and the resolution of the ellipsis. According to the alternative, "lexical affixation" account, affixation of CMPR directly to *proud* means that the ellipsis has to be resolved as in (26), if standard conditions on ellipsis are assumed:

(26) John is [prouder of Bill] than Mary is <prouder of Bill>.

This is not the correct result. The question is how the Lexicalist account can force the Deg element *not* to be interpreted in the ellipsis position, given that the rest of the word that it is attached to (*proud* above) is interpreted there. It would be possible to stipulate a solution this problem: it could be claimed that degree elements do not count for the calculation of ellipsis conditions, or that they do not figure in the resolution of ellipsis. Such an account would have to specify what it would mean for a subpart of a word—with the word an atom as far as the syntax is concerned—to be invisible for the purposes of ellipsis resolution. In any case, further considerations show that this attempt at a fix is itself problematic. Consider (27):

(27) Bill is prouder of Mary's longest book than Fred is Δ.

Here Δ is resolved to <proud of Mary's longest book>. If degree elements are systematically ignored in ellipsis, then the only possible interpretation of should be *Bill is prouder of Mary's longest book than Fred is proud of Mary's long book*. This is not the case.

The syntactic account suffers from none of these problems. It is able to account for the scopal behavior and the ellipsis cases uniformly, with the assumption that Deg is a separate syntactic element. I take this to be evidence in favor of the syntactic approach that I am pursuing here.[16]

[16] While this argument is stated in terms of ellipsis in the text, it is possible that the same point can be made in terms of approaches that do not have ellipsis in the relevant examples. To take a case in point, Kennedy (2002) follows earlier work in proposing representations like the following:

3.2 Metalinguistic comparatives

The treatment above involves a uniform syntactic analysis of Deg in conjunction with (prosodically conditioned) Local Dislocation to account for the morphophonological forms of comparatives and superlatives. This analysis must take into account cases of *metalinguistic comparison*, in which *more* appears adjacent to adjectives that normally take a synthetic form. It has been observed in the literature that cases of metalinguistic comparison provide some insight into the conditions on synthetic comparative formation (see, for example, di Sciullo & Williams, 1987), by showing that surface adjacency is in some sense not enough for the creation of a synthetic form:

(28) Metalinguistic Comparisons
  a. John is more sad than tired.
  b. John is sad more than tired.
  c. *John is sadder than tired.

As shown in (28b), *more* can appear post-adjectivally with the metalinguistic comparative; this is not possible with normal comparatives; for clarity, I refer to examples like *John is taller than Bill* as "true" or "normal" comparatives. Because *more* and an adjective can surface as linearly adjacent in metalinguistic comparatives, at a minimum it must be the case that surface linear adjacency is a necessary but not sufficient condition on the Local Dislocation process. The examination of such comparatives below shows (1) that, following earlier accounts, metalinguistic comparatives differ syntactically from normal comparatives; and (2) that the structural difference results in Deg and the adjective not being adjacent in the way that is required for the formation of synthetic comparatives.

3.2.1 Initial points

Although observations about metalinguistic comparative formation are to be found, little has been said about the exact reason why synthetic comparatives like (28c) are impossible. In this subsection, I develop an account of the relationship between the syntax of metalinguistic comparatives and the morphosyntax of analytic and synthetic forms. The main thrust of this argument is that an account of the prohibition of synthetic forms in the relevant cases can be given in structural terms, in a manner that follows rather closely Bresnan's (1973) discussion of metalinguistic comparison. In particular, it can be demonstrated that Deg and the adjective are not adjacent in metalinguistic comparatives in a way that could trigger the rule of Local Dislocation formulated above. The absence of synthetic comparative forms is thus predicted directly from the structural analysis.

Before the morphophonological matters are addressed, some points about metalinguistic comparison must be clarified, since there is apparently some uncertainty

Footnote 16 continued

(i) John is [$_{DegP}$ er [$_{AP}$ proud of Bill] than [$_{CP}$[$_{DegP}$Ø [$_{AP}$ p~~roud of Bill~~]] Mary is [$_{DegP}$Ø [$_{AP}$ ~~proud of Bill~~]]]]

The mechanism involved here is not ellipsis, but nevertheless requires identity. If *prouder* were formed lexically in the matrix AP, it is not clear that the necessary identity conditions would be met, in which case deletion should be impossible. This type of analysis would therefore not work for reasons parallel to those discussed in the text, unless the conditions on identity were specified differently.

concerning the scope of this phenomenon. In McCawley (1988:673), the following examples are given, presented here with the judgments that he assigns:

(29)a. Your problems are more financial than legal.
b. *Your problems are more financial than mine.

Here (29a) is a metalinguistic comparative, one that means that it is more appropriate to call your problems financial than it is appropriate to call them legal. The choice of *financial* here is an attempt to force a metalinguistic comparison, which, unlike a normal comparative, does not have to occur with something gradable. The example (29b) is apparently supposed to be deviant on the grounds that it could not be metalinguistic; as a result, it is interpreted with a normal comparative Deg and the adjective *financial*, and this is problematic because of the properties of this adjective. Note that this analysis relies on the idea that (29b) is not metalinguistic comparison because it cannot be paraphrased like (29a) ("your problems are more (what I would call) financial than (what I would call) legal"). The other assumption seems to be that the metalinguistic comparison is only possible when the *than*-clause contains something identical in category to the element that appears after (or before) *more*; see below.

The problem with this initial characterization—whether this is an accurate depiction of the assupmptions underlying McCawley's discussion or not—is that (29b) is grammatical on the interpretation 'your problems are more appropriately called financial than mine are appropriately called financial'.[17]

For reasons that are made clear immediately below, I will consider the examples above in connection with examples of the following type, which could conceivably be a different type of comparative; see below:

(30) Fred's hair is more short than John's is.

This example might seem somewhat deviant out of context, like many comparatives of this type. However, it has an interpretation which seems to be quite similar to what is found in the *more hot than humid* type examples. An initial attempt to specify what is happening in comparatives like (30) appears in (31):

(31) John is more smart than Bill is.
a. Assertion: It is more appropriate to call John *smart* than it is to call Bill *smart*.
b. Implicature: John is smarter than Bill.

It should follow then that the apparent deviance of e.g. *John is more smart than Bill* out of the blue stems from the fact that the assertion is of the highly specific kind sketched above (not to mention the relationship of the implicature to the assertion of the 'true' comparative).

In the appropriate context, the implicature can be cancelled. Consider a situation in which Mary is 8 cm taller than the average height for females, whereas John is 4 cm taller than the average height for males. Suppose in addition that in absolute terms, John is taller than Mary. In such a context it is possible to assert (32):

(32) Mary is more tall than John is (but of course John is taller in absolute terms).

[17] It is also possible, although perhaps more difficult, to get the reading "Your problems are more appropriately called *financial* than appropriately called *mine*".

In the example (32) the assertion is as above: it is more appropriate to call Mary tall than it is to call John tall. Because what it means for Mary to be appropriately called tall is different from what it would mean for John to be called the same, it is possible for John to be taller than Mary in absolute terms. The implicature derived from (32) is that Mary is actually taller than John in absolute terms, but this can be cancelled, either by the context or by additional linguistic material.

Alternatively, it could be said that in the normal case, the scales for *tall* with respect to *John* and *Mary* are identical, from which the implicature to the normal comparative would follow. In (32), however, the standards for the scales can be interpreted as different (e.g. "tall for a male"; "tall for a female"), with the difference in scales being compatible with John being taller in the absolute sense.[18]

As expected, the effects of cancellation are not found with the normal comparative, which has a synthetic form:

(33) #Mary is taller than John is (but of course John is taller in absolute terms).

The reading identified for comparatives like (31) above is not restricted to adjectives that take synthetic forms; any other predicate should have the same 'meta' reading as well, with acceptability depending to some extent on the properties of the predicate.[19]

The type of metalinguistic comparative that has been noted more frequently in the literature involves a case in which the applicability of two adjectives is at issue; for instance:

(34) John is more lazy than stupid.

Here the assertion is the same as that found in the cases examined immediately above:

(35) Assertion: The term *lazy* is more appropriately applied to John than the term *stupid* is.

The existence of an implicature that parallels what is seen in the examples above is difficult to determine for certain choices of what is being metalinguistically compared. For examples like (34), the reason for this is if it were parallel, it would be something like 'The degree of John's laziness is greater than the degree of his stupidity'. The problem is that this is difficult to assess when there are incommensurable scales. This is easy for cases like *This table is more wide than it is tall* versus *This table is wider than it is tall* but difficult where there is no obvious scale for true comparison; hence the strangeness of *John is lazier than he is stupid*. The latter type of example is acceptable to the extent that the scales associated with laziness and stupidity can be made commensurable. Since this is difficult, the sentence is deviant accordingly.

[18] It might very well be that what we are dealing with here is a "Comparative of Deviation", although there are some questions about how this might work; see below. As a reviewer notes, the treatment of this latter type as a comparative does seem to get the semantics right.

[19] Here 'predicate' is used in an informal sense. Consider e.g. *John is more Las Vegas than Monte Carlo*, where the metalinguistic reading arises with two (proper) nouns. There are evidently some restrictions on how high metalinguistic comparison can be; it cannot, for instance, attach to clauses (TP?):

(i) *More John is incompetent than Bill is lazy.

The intended comparison in (i) is like that found in *It is more the case that John is incompetent than it is the case that Bill is lazy*.

An anonymous reviewer suggests that (30) and (32) are instances of what Kennedy (1997, 2001) calls a *comparative of deviation* (COD). I am entirely sympathetic with this suggestion, since my primary concern is morphophonological and metalinguistic comparatives and CODs both disallow synthetic adjectives. At the same time, there are some points that must be sorted out in this particular classification, which I note before moving on.

According to Kennedy, the COD is a comparatives in which "...the relative extents to which two objects deviate from some standard value associated with the adjective" are compared (Kennedy, 1997:252). Kennedy's COD examples all involve pairs of polar adjectives, like (36).

(36) San Francisco Bay is more shallow than Monterey Bay is deep.

Kennedy notes that CODs prevent the formation of synthetic comparatives, noting that this could be because they differ structurally from normal comparatives. In light of the discussion above, the point about the COD leads to the overall suggestion that there are two types of non-standard comparatives at play:

(37)a. Metalinguistic Comparative: more X than Y, where X and Y are of the same category. Example: John is more lazy than stupid.
b. Comparative of Deviation: A is more X than B is (Y).

The idea is that examples like *John's hair is more short than Fred's (is)* are instances of COD, and not metalinguistic comparatives. If this classification is correct, then (38a) must be a metalinguistic comparative, whereas (38b) must be a COD:

(38)a. John is more lazy than stupid.
b. John is more lazy than Bill is stupid.

If (38b) is a COD, then, unlike what is found in metalinguistic comparatives, there must be commensurable scales. According to the reviewer mentioned above, (38b) is as deviant as *??John is lazier than he is stupid* is; in the latter type of case, the deviance results from incommensurable scales. This, however, is not my judgment, nor that of other speakers I have consulted, for (38b). The sentence (38b) receives an interpretation like "John is more appropriately called lazy than Bill is appropriately called stupid"; i.e., that of the (38a) type cases.

How these conflicting intuitions are to be reconciled—i.e. how to untangle the potential differences between "true" metalinguistic comparatives and CODs—is not clear. It might be that the two cases in (37) are quite similar. Some of this is clear from the interpretation. Kennedy (2001) paraphrases COD's as follows:

(39)a. The Red Sox are more legitimate than the Orioles are fraudulent.
b. Paraphrase: The degree to which the Red Sox exceed a standard of legitimacy is greater than the degree to which the Orioles exceed a standard of fraudulence.

For e.g. *John is more tall than Bill is*, the paraphrase in terms of a COD would therefore evidently be as follows:

(40) Paraphrase: The degree to which John exceeds a standard of tallness is greater than the degree to which Bill exceeds a standard of tallness.

This is similar in many ways to the paraphrases in terms of appropriateness that I have offered for metalinguistic comparatives, especially to the extent that the determination of what is a standard value relates to the notion of appropriateness that has been invoked above. It could be that both readings are in principle available, and involve differences in what is being compared: appropriateness versus deviation. This is touched on below.

For the study of analytic/synthetic alternations, the important point is that for the purposes of morphophonology, both types of comparative in (37) seem to behave identically. The same is true of the syntax; both, for instance, allow post-predicate *more*:

(41)a. John is lazy more than stupid. (metalinguistic)
b. John is lazy more than Bill is (stupid). ("COD")

And, as noted, Kennedy (1997, 2001) suggests that COD constructions have different syntactic properties than normal comparatives do, properties which make them, like the metalinguistic comparatives, prevent the formation of synthetic comparative adjectives. For these reasons, I treat them in the same way below, using "metalinguistic" as a cover term.

3.2.2 Structure

What is primarily at issue is the structure of the metalinguistic comparatives viewed as a whole, and the specific question of why they disallow the formation of synthetic comparatives. Bresnan (1973:325ff.) proposes that the comparative element in examples of the type *I'm more sad than I am happy* is not part of the adjective phrase to begin with; rather, it is in a phrase that is adjoined to a higher projection. The basic components of this analysis are adopted here.

One major difference between normal and metalinguistic comparatives is that the metalinguistic type allow post-adjective *more*:

(42) John is lazy [more than stupid].

This is never possible with true comparison, which is to say that examples like *John is obnoxious more than Mary* can only be interpreted as metalinguistic comparatives.[20]

Illustrating with *John is more lazy than stupid*, Bresnan's structure is essentially that in (43), where I have attached the DegP *more than stupid* to the AP:[21]

[20] Note that in addition to metalinguistic comparison, there is a related structure for (42) with an interpretation that relates to times; this is clearest when two stage-level adjectives are involved:

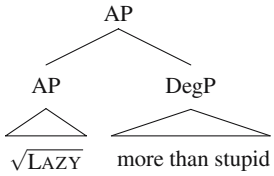
(i)a. John is drunk more than he is stoned.
b. =more often drunk than stoned (or drunk more of the time).

I put these temporal readings aside in the discussion below.

[21] This is based on the fact that this entire phrase behaves like an AP constituent:

(i)a. [Lazy more than stupid] though John may be...
b. [More lazy than stupid] though John may be...

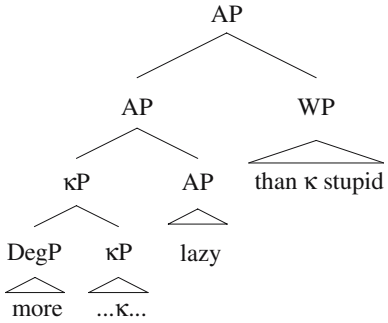
(43) Bresnan's Structure (adapted)



I assume that (43) is the gross structure for *John is lazy more than stupid*. In the pre-adjectival *more* cases, the DegP of metalinguistic comparison is linearized on the left, and whatever operation forces the *than-* clause to appear after the adjective in normal comparatives applies.

Along with these structural properties, metalinguistic comparatives (and CODs) have distinctive interpretive properties that differ greatly from normal comparatives. Most strikingly, the degree that is being compared is one of appropriateness (or deviation), not a degree associated with the adjective itself. For this reason I assume that metalinguistic comparatives involve a silent adverbial element—given here as κ heading κP—that provides the semantic properties of "appropriateness" that are found in comparatives of this type. DegP is attached to κP in metalinguistic comparatives, and to AP in normal comparatives. Thus for the derivation of examples like *John is more lazy than stupid*, this approach holds that κP originates as an adjunct on the AP headed by *lazy* (in the post-adjectival case noted above, κP is linearized on the right):

(44) Structure of Metalinguistic Comparative



Although a number of points remain to be specified, the insight behind this analysis is that the *more* of metalinguistic comparison is found in a phrase that combines with the adjectival structure in very different way than the DegP containing *more* in true comparatives. The same is true of the CODs discussed above. In structural terms, CODs are evidently identical to metalinguistic comparatives with κ. It could be that the difference between these two types is found in the fact that in CODs, it is not κ for 'appropriateness' but η for 'deviation' that is the silent adverbial.[22]

[22] Above I noted the existence of a temporal reading for certain examples like *John is drunk more than stoned*. In this case, κ does not appear; instead, the null element is an adverbial τ. As a point

Neither metalinguistic comparatives, CODs, nor comparatives with the temporal reading (fn. 22) allow for the formation of synthetic comparatives. This fact is explained when the details of linearization and Local Dislocation are extended to the structures examined above.

3.3 Structure and linearization

The syntactic and semantic discussion of the last subsection is prompted by the absence of synthetic forms in metalinguistic comparatives, something that calls for an explanation on an adjacency-based view. These comparatives constitute instances in which Deg is linearly adjacent to an adjective which could in principle take a synthetic form, but Local Dislocation does not apply. One way of looking at what is at issue here is in terms of a distinction between 'principled' and 'accidental' linear adjacency. This type of statement can be made directly in approaches which posit (sets of) mapping relationships between different levels of linguistic representation. For instance, it would be possible to distinguish directly between adjacency as a linear relationship that expresses some other grammatical relationship on the one hand from accidental linear adjacency on the other. The equivalent of Local Dislocation might then be specified to apply only in the former case, but not the latter. See, for instance, the theory of Marantz (1984), particulary p. 287ff. In an approach like that assumed here a statement of this type cannot be made directly. Rather, the differences between two cases of surface adjacency must be the result of either (1) the structures involved; or (2) the order of operations (e.g. phase-based spell-out). The absence of synthetic comparatives under adjacency in metalinguistic comparatives is plausibly the result of either, assuming the structural analysis above.

In order to make this point explicit some more details concerning linear relations are required. For the purposes of this discussion, attention is restricted to relationships between heads and phrases. Since the crucial issue is how the Deg head is linearized with respect to other heads, head-internal structure is not relevant, because the issue hinges on how the M-Word Deg combines with the M-Word adjective (some details concerning the relationship between M-Word-internal and -external structure appear in Embick (2005)).

The approach to linearization that is developed here follows Sproat (1985) and related work, where linearization is treated as a two-step process. The first step involves addition of the *-operator, in an algorithm that targets each node at the M-Word level and higher in terms of phrase-structure status; this establishes sets of adjacency statements that refer to heads (M-Words) and phrases (here [X Y] is used to refer to the node that has X,Y as daughters):

(45) LIN[X Y] → (X * Y) or (Y * X)

Footnote 22 continued
concerning relative syntactic distribution, the κP with the meta-reading is associated with a lower attachment site than τP:

(i)a. John is more drunk than stoned more than Bill is
b. John is drunk [more than stoned] [more than Bill is]
c. *John is drunk [more than Bill is] [more than stoned]

The example (ic) is ungrammatical on the relevant reading (that of (ia–b)). Rather, (ic) means only that John is drunk to a greater degree than Bill is more frequently than he is stoned.

The statement (X * Y) says that X has to be left-adjacent to Y. In cases in which X or Y is phrasal, something more must be said about how the individual terminals come to be concatenated. The reason for this is that the phrase is abstract, unlike the terminals. The specification of a linear order requires that the terminal nodes be linearized (concatenated). I assume that what it means for head X to be left-adjacent to YP is for X to be left-adjacent to the left-most element of YP.[23] This second step is a process of concatenation in which the set of relationships defined by * is augmented by statements in terms of the concatenation operator $^\frown$. The intuition here is that * specifies what is next to what in relatively abstract terms, which relate directly to the hierarchical properties of the phrase marker. For instance, (V * DP) for e.g. English manifests the fact that this language is head-initial; what this means is that V occurs to the left of the DP, whatever that DP may happen to contain. When it comes to the concatenation of terminal nodes, this means that V is concatenated with the first element of the DP. This is where the concatenation step comes into play.

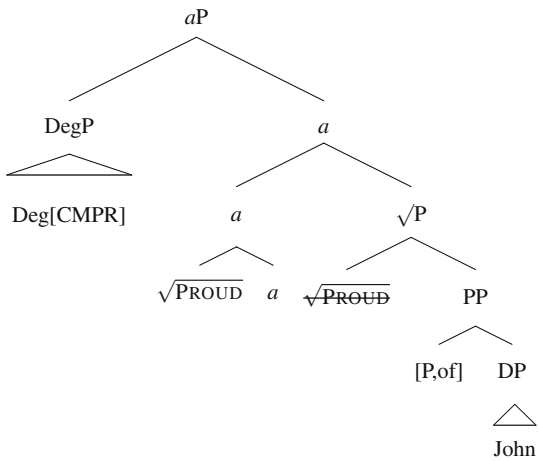
For present purposes, the effects of the concatenation process are stated in (46):

(46) For $X(P) = [W_1 \ldots W_n]$ and $Y(P) = [K_1 \ldots K_n]$, where W_i, K_i are M-Words, $(X(P) * Y(P)) \rightarrow (W_n {}^\frown K_1)$

There is much more to say about (46) and how it operates; at this point, it expresses the effects of what an explicit procedure should do, but is not itself such a procedure.

For an illustration of the process sketched above consider *prouder of John*, which has the structure in (47) (recall that in this structure the Root $\sqrt{\text{PROUD}}$ moves to *a*)):

(47) Structure



The LIN procedure applies to each node X in this structure to yield a set of additional statements with *, which specify a relationship of adjacency between the

[23] This is employed in a related fashion in Marantz's (1984) discussion of cliticization. A similar statement is employed in Fox and Pesetsky (2005).

two daughters of X. In order to be explicit about the structure and the linearization procedure, I have expanded "A" to include *a* and the $\sqrt{\textsc{Root}}$ in these statements:[24]

(48) Addition of *
LIN[[P,of] DP] → ([P,of] * DP)
LIN[$\cancel{\sqrt{\textsc{Proud}}}$ PP] → ($\cancel{\sqrt{\textsc{Proud}}}$ * PP)
LIN[*a* $\sqrt{}$P] → ([$\sqrt{\textsc{Proud}} \oplus a$] * $\sqrt{}$P)
LIN[DegP *a*] → (DegP * *a*)

As noted, general definitions concerning headedness and so on could be invoked in justifying specific outputs; e.g. the fact that English is head-initial should determine (P * DP) as opposed to (DP * P), and so on.[25]

All other things being equal, it is assumed that these statements are additive; i.e., they are statements added to the representation in addition to the set of statements that define the hierarchical properties of this phrase marker. As noted above, in addition to what appears in (48) there must be additional statements for the subparts of M-Words, the Subwords; the statements above define the ordering of M-Words and phrases.[26]

The second step, associated with (46), imposes ⌒ in the way described above; continuing with the example based on (47), this additional set of statements is as follows:[27]

(49) Addition of ⌒
([P,of] * DP) ⟶ [P,of]⌒John
([$\cancel{\sqrt{\textsc{Proud}}}$ * PP) ⟶ $\cancel{\sqrt{\textsc{Proud}}}$⌒[P,of]
([$\sqrt{\textsc{Proud}} \oplus a$] * $\sqrt{}$P) ⟶ ($\sqrt{\textsc{Proud}} \oplus a$)⌒ $\cancel{\sqrt{\textsc{Proud}}}$
(DegP * *a*) ⟶ Deg[CMPR]⌒($\sqrt{\textsc{Proud}} \oplus a$)

With string adjacency at the M-Word level established by the statements in (49), the rule for synthetic comparative formation is stated in terms of ⌒, as in (50); this rule replaces the initial formulation of C/S Local Dislocation formulated in (11) above, which was stated in terms of *:

(50) English C/S Local Dislocation (revised)
Deg[CMPR,SUP] ⌒ $[...X...]_a$ ⟶ $[...X...]_a \oplus$Deg[CMPR,SUP]
where the phonological form of $[...X...]_a$ meets the prosodic condition

More generally, the hypothesis is that all Local Dislocation operations are defined in terms of statements derived by the normal linearization mechanisms: in particular, in terms of concatenation statements like (X⌒Y). See Embick (2005) for some discussion.

[24] The '[P,of]' here is shorthand for the preposition head P that receives the phonology of *of*. The internal structure of the DP containing *John* is also simplified, since I am not taking a stance about the syntax of proper names.

[25] Something further must be said for apparent cases in which such generalizations about headedness do not apply; see Kroch (1994) for examples of this type.

[26] It is also the case that something in the system must distinguish e.g. the non-M-Word projection of *a* that dominates the M-Word *a* and $\sqrt{}$P from the M-Word *a* (itself internally complex). The same issue is touched on in a related guise in Richards (2002).

[27] There must also be a procedure for determining which copy of an element is pronounced.

For the derivation of *prouder than John*, there is a statement in the last line of (49) that meets the environment for (50). The effects of the application of (50) are shown in (51):[28]

(51) $\text{Deg[CMPR]}^\frown(\sqrt{\textsc{Proud}} \oplus a) \longrightarrow ((\sqrt{\textsc{Proud}} \oplus a) \oplus \text{Deg[CMPR]})$

In this way, the rule (50) transforms the ordering statement '$\text{Deg[CMPR]}^\frown(\sqrt{\textsc{Proud}} \oplus a)$' on the left-hand side into the representation on the right, where Deg is a Subword affixed to a.

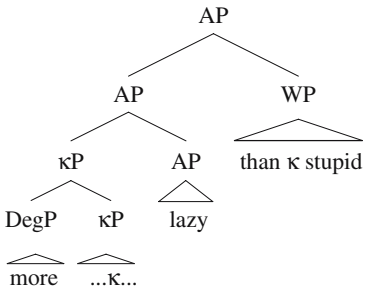
A final question concerns the statement that is the output of the Local Dislocation rule, and how it relates to other concatenation statements. The complex head after the Local Dislocation—i.e. the comparative adjective $((\sqrt{\textsc{Proud}} \oplus a) \oplus \text{Deg[CMPR]})$—has to appear immediately left-adjacent to $\cancel{\sqrt{\textsc{Proud}}}$. The statements in (49) specify an ordering on the adjective prior to the application of (51). The same relative ordering between $[\sqrt{\textsc{Proud}} \oplus a]$ and $\cancel{\sqrt{\textsc{Proud}}}$ must be maintained after Deg has been affixed as a Subword to the adjective. One way to accomplish this is with reference to what the statements in (49) contain in the first place. While in (49) the Subword components of individual M-Words are represented, the rule (46) that introduces $^\frown$ is defined at the M-Word level; i.e., it orders M-Words with respect to other M-Words directly, not by virtue of their subparts. The internal structure of the M-Words is irrelevant to these ordering statements; as long as the individual M-Words can be distinguished from one another, no crucial reference to their internal structure must be made in the statements derived by (46). The upshot of this is that the adjective has the same linearization status after affixation of Deg as it does before; that is, when Local Dislocation creates $Y \oplus X$ from $(X * Y)$, the derived object $Y \oplus X$ inherits or retains the linearization contraints imposed on Y (cf. also Marantz (1984, 1988)). Perhaps it has the requirements of X as well, although this is only one possibility.[29]

The rules above account for synthetic C/S-formation in the case of normal comparatives. In the case of metalinguistic comparatives, there are two observations to be made. First, concerning linearization statements like those defined in terms of $^\frown$ above, the Deg head is never in a local relationship with the adjective, so that (50) is not triggered. Recall the structure for *John is more lazy than stupid*, repeated from above:

[28] For present purposes it is assumed that the bracketing of Deg[CMPR] outside of $(\sqrt{\textsc{Proud}} \oplus a)$ follows from a general account of how an M-Word is converted into a subpart of another M-Word by Local Dislocation.

[29] The point about the requirements of X relates to a further question concerning material to the *left* of Deg[CMPR]. For instance, in *John is taller than Bill*, the possibilitiy exists that there is a statement ordering a copy of *is* with respect to Deg[CMPR]: $[v, \cancel{\text{is}}]^\frown\text{Deg[CMPR]}$. If this is the case, then the object derived from Local Dislocation must have the properties of both of its constituent elements, so that the proper ordering between elements left of Deg[CMPR]'s original place of linearization and the derived synthetic comparative are maintained. This can be specified in terms of conjunction. Alternatively, it is possible if linearization operations are strongly cyclic that ~~is~~ is only linearized with respect to the synthetic adjective in the first place, i.e. after Local Dislocation has taken place. I assume that further investigation will reveal which of these options is correct, since examining the issue in greater detail is not possible here.

(52) Structure of Metalinguistic Comparative



Given the rules above, the linearization procedure applied to this structure does not generate the linearization statement in (53), assuming that DegP is linearized to the left of κP, just like it is with normal APs:

(53) DEG[CMPR]$^\frown$[$_A$LAZY]

The reasons for this are structural, since the linearization process applied to the tree in (52) cannot create this statement. Only (DEG[CMPR]$^\frown$ κ) is generated. In the absence of (53) the Local Dislocation rule does not apply, and [CMPR] is supported by *mo-* just as it is in other cases in which it is not affected by Local Dislocation.

This first explanation relies on the idea that null elements like κ are counted in concatenation statements. In the case at hand, there is potentially a second reason for the absence of synthetic metalinguistic comparatives. Assuming that linearization and spell-out to phonology occur in phases, and that κP is a phase, Deg is supported by *mo-* at a stage in the derivation before it comes to be linearly adjacent to an adjective. On the assumption that κP in (52) is adverbial/adjectival, the idea behind cyclic spell-out is that the phonology of Deg has been taken care of inside of κP before Deg comes to be in any sort of relationship with the AP containing $\sqrt{\text{LAZY}}$. Inside of κP the rules for analytic forms apply, since there is no Local Dislocation (the adjective is not present at this derivational stage), such that *more/most* surfaces. This account of why synthetic comparative formation does not occur requires specific assumptions about how phases are defined in terms of category-defining projections, and, in addition, some assumptions about how adverbial-like modifiers fit into this system. That is, different theories of phases make different predictions about when the DegP should have been processed, and these differences are of course relevant for this analysis.

Distinguishing between the two explanations advanced above might be possible when other case studies are examined, but I will not attempt to make such a distinction here.

To summarize, when the structural properties of metalinguistic comparatives are examined closely, the absence of synthetic comparative forms can be explained. In this particular case there are in fact two coherent explanations for why the Local Dislocation rule above does not apply in such cases, each of which leads to further questions to be investigated empirically. The solutions are stated in terms of explicit assumptions about how linear order is imposed on syntactic structures. A number of additional questions concern how these linearization operations are interleaved with other operations, particularly given further assumptions currently under discussion in the literature, e.g. the idea that the construction of PF proceeds in parallel to the syntactic derivation. In addition, much more could be said about movement and

the status of unpronounced copies in such a system; some questions of this type are studied in Fox and Pesetsky (2005).[30]

3.4 The adverb–adjective cases

The metalinguistic comparatives discussed above are an instance of potential surface adjacency without the formation of a synthetic form. As demonstrated above, this is not a problem for a view based on structure and adjacency like the one I am advancing here. Another case that requires further study involves the premodification of adjectives by adverbs, as in *Mary is more amazingly smart than John*. These cases seem to show that the creation of synthetic forms is prevented when an element appears linearly between Deg and the adjective, along the lines predicted by the syntactic account. Further points concerning these cases remain to be clarified, however.

As an initial step in the description, there are roughly speaking two types of cases to consider: one in which the adverbial is evaluative, and one in which it specifies the 'manner in which' the subject of predication is ADJ:

(54) Roughly evaluative (Type 1)	(55) "Manner in which" (Type 2)
amazingly smart	ploddingly slow
incredibly tough	happily drunk
unbelievably short	coyly glib
surprisingly long	flatly honest
annoyingly nice	venomously sarcastic
frightfully bold	brutally honest
pleasingly sweet	rudely late

There is a paraphrase relationship that differentiates these two types, as shown in (56–57); the paraphrases highlight the fact that Type 1 seem to be related to a degree in a way that does not involve the Deg head; Type 2 adverbs do not have this property:

(56) Type 1
a. John is incredibly intelligent.
b. The degree to which John is intelligent is incredible.

(57) a. John is ploddingly slow.
b. ≠The degree to which John is slow is plodding.

[30] Some other questions concern the ordering of operations and the effects of this on allomorphy and other phenomena. For instance, it could be asked if morphemes with null (-Ø) phonological exponents are 'invisible' for the purposes of the factors conditioning allomorphy, and how this relates to statements about linear order. Some preliminary considerations related to this point are advanced in Embick (2003).

It is conceivable that a solution to the 'bracketing paradox' posed by *unhappier* could be phrased in these terms. If Vocabulary Insertion at NEG occurs after $\frown$ relationships have been defined, there is no paradox.

(i) $\text{DEG}\frown(\text{NEG}\oplus(\sqrt{\textsc{Happy}}\oplus[a,\emptyset]))$

As far as the rule (50) is concerned, the Local Dislocation applies; the target *a* meets the relevant phonological condition, because the phonology *un-* of NEG has not been added to this node (perhaps because this object does not define a spell-out domain). A similar perspective is advanced by Newell (2004), who employs 'late adjunction' to a similar end. See also Speyer (2005) for a structural approach to these paradoxes.

To a first approximation, the first type asserts something evaluative about a degree associated with the adjective, while the second type asserts something about the manner in which the relevant entity is A, so that for instance the *ploddingly slow* are a subset of the *slow* who are slow in that particular fashion.[31]

In terms of the linearization mechanics of the preceding subsection, all of these instances of [Deg [ADV ADJ]] are linearized to produce Deg⌒ADV.[32] It is correctly predicted by the view of C/S-formation advanced here that only analytic forms should be possible in cases of this type. By illustrating a type of intervention effect in an object that is constructed syntactically, these cases fit nicely with the syntactic approach to Deg. However, there are in the literature two challenges to the view that these cases are informative for the study of analytic/synthetic alternations. One is the position of Williams (2004), who suggests that [Deg [ADV ADJ]] is always metalinguistic. The second is found in Kiparsky (2005), who suggests that Deg only operates on the adverb in e.g. *more amazingly smart*. These are addressed in turn.

3.4.1 Only metalinguistic?

An initial question for a more detailed analysis of [ADV ADJ] comparatives and superlatives is whether they are capable of being normal comparatives in the first place. Williams (2004) suggests that examples like *John is the most amazingly smart student* are uniformly interpreted as a kind of metalinguistic comparative. If this were true, then the [ADV ADJ] examples would show nothing about C/S-formation beyond

[31] Strictly speaking things like *amazingly* can also have a higher position, e.g. *John is, amazingly, smarter than Bill means It is amazing that John is smarter than Bill.*

(i)a Type 1
i. John is more incredibly intelligent than Bill.
ii. is incredibly more intelligent than Bill

b. Type 2
i. Fred is more ploddingly pedantic than Larry.
ii. *Fred is ploddingly more pedantic than Larry.

With some of the adverbs that appear in (55), it is possible to have the higher scopal position as well (e.g. *happily*).

[32] While Type 1 and Type 2 adverbs behave similarly in not appearing after the adjective, there is a third, superficially similar type of which the same is not true. To a first approximation, cases like *physically strong* appear similar to Type 2 above; but consider (i):

(i) Mary is physically stronger than John.

Based on the fact that the synthetic comparative can be formed here, it must be the case that *physically* attaches higher than the adverbs in the examples above (otherwise it would intervene linearly between Deg and the adjective). Correspondingly, adverbs of this type can appear post-adjectivally, unlike the others examined above:

(iii)a. Mary is strong physically.
b. i. *Susan is smart amazingly.
ii. *John is insane dangerously.
iii. *Bill is slow ploddingly.

This shows that despite superficial similarities, *physically strong* is different structurally from *ploddingly slow* (and, for that matter, from *amazingly smart*); the natural way to think of this would be to take the adverbials of the *physically* type to attach to a different part of the structure than the Types 1 and 2 adverbs do, one that allows different linearization possibilities; these cases are not considered further here, although they do raise some questions.

what can be learned from the metalinguistic cases discussed in the last section. Part of Williams' proposal is certainly correct: *John is more amazingly smart than Bill can* be interpreted as a metalinguistic comparative, in the same way that all of the other analytic comparatives examined to this point can be. The question is whether the [ADV ADJ] cases *must* have this interpretation and only this interpretation (i.e. whether it is true that they never have 'normal' comparative readings).

This question is complicated. The diagnostics discussed above show when a particular structure must be interpreted as a metalinguistic comparative. For example, if a comparative shows a post-adjectival *more*, then it must be a metalinguistic comparative. Or, when a normally synthetic comparative surfaces as analytic, the interpretation is metalinguistic. None of the diagnostics employed to this point go in the other direction, showing that a particular case must be interpreted as a 'true' comparative and not a metalinguistic comparative.

One way of approaching the question involves the particular conditions that are associated with metalinguistic readings. Out of context, many examples that have only metalinguistic readings are quite difficult to interpret. Consider, for instance, comparatives of adjectives modified by *very*:

(58) John is more very smart than Bill is.

Examples like (58) are highly deviant out of context; they are possible only as metalinguistic comparatives, in which *very smart* has a special prosodic status associated with a particular intonational contour (like having quotes around it):

(59) John is more || very smart || than Bill is.

This does not seem to be the case with e.g. *amazingly smart* and other examples like those above in comparative contexts, although different intonational breaks do in fact seem to correlate systematically with different bracketings (see Sect. 3.4.2). With *amazingly smart*, prosodic breaks like in (59) are possible and when present force the metalinguistic interpretation. But the fact that such obvious breaks are not required with *more amazingly smart* and the like suggests that the structure for metalinguistic comparatives is not the only one possible for such strings.

There appear to be more definitive diagnostics showing that a 'normal' comparative reading is available for certain cases. One set of examples is based on comparatives that appear pre-nominally. In such examples, regular comparison is possible:

(60)a. [A smarter person than Mary] is difficult to find.
b. [A more intelligent person than Mary] is difficult to find.

On the other hand, when metalinguistic readings are forced in this context by using an analytic form for an adjective that normally forms synthetic comparatives, the result is deviant:

(61) ?*[A more smart person than Mary] is difficult to find.

The comparatives of [ADV ADJ] cases do not have the status of (61), suggesting that, all other things being equal, these are cases of true comparison:

(62)a. [A more amazingly smart person than Mary] is difficult to find.
b. [A more ploddingly slow person than John] is hard to imagine.

The example in (63) provides a baseline for another argument that the [ADV ADJ] cases in comparison do not behave only as metalinguistic comparatives:

(63) John wants to be taller than Bill Δ.

In examples of this type there is an ambiguity depending on the resolution of the ellipsis:

(64) Readings of (63); Δ =
a. <wants to be δ-tall>
b. <is δ-tall>.

For reasons having to do with the English tense system and auxiliary verbs, the same ambiguity is not found in (65):

(65) John wants to be taller than Bill does Δ.

Examples of this type allow the ellipsis to be interpreted only as wide, the reading in (64a).

The relevance of these ellipsis patterns is that 'true' comparatives and metalinguistic comparatives can be distinguished in these terms; in particular, true comparatives display the ambiguities seen in (63), while metalinguistic comparatives do not. An example is provided in (66):

(66) John wants to be more tall than Bill Δ.

In terms of the silent κ used above for metalinguistic comparatives, the ellipsis can only be resolved as in (67):

(67) $\Delta = <$ is κ tall $>$

I.e.: John wants to be more κ-tall than Bill is κ-tall

This is the narrow ellipsis; it seems that the wide ellipsis is unavailable in examples of this type. This impression is confirmed by (68), which adds *does* to (66).

(68) *John wants to be more tall than Bill does Δ.

The use of the *do* forces the wide ellipsis, which is absent in cases of metalinguistic comparison. As a result of this conflict, (68) is ungrammatical.

A similar effect is found with *claim*:

(69)a. John claims to be more tall than Bill Δ.
$\Delta = <$ is κ "tall">
b. *John claims to be more tall than Bill does Δ.

If these considerations are on the right track, then something like (70) holds:[33]

(70) Metalinguistic comparatives in certain contexts allow only narrow ellipsis resolution.

The importance of (70) for my purposes is that it can be applied to the *amazingly smart* cases. If *John is more amazingly smart than Bill* had only a metalinguistic comparative interpretation, then it should allow only narrow ellipsis in the relevant context. The relevant context is given in (71), with the possible resolutions of the ellipsis:

[33] For my purposes it suffices that (70) holds, but naturally one would like to know why. A reviewer suggests that (70) holds because the result of scoping a metalinguistic comparative out of its clause results in deviance.

(71) John wants to be more amazingly smart than Bill Δ.
a. <wants to be δ-[amazingly smart]>
b. <is δ-[amazingly smart]>

The wide reading is available in (71); thus according to the reasoning above, it must be the case that *more amazingly smart* can have the reading of a true comparative. This is confirmed by an example with *does*:

(72) John wants to be more amazingly smart than Bill does Δ.

There is a contrast between cases of this type and examples like (68) above. The acceptability of (72) again suggests that *more amazingly smart* is a true comparative, as it patterns with those cases and not the cases of metalinguistic comparison.

As an interim conclusion, this line of argumentation shows that the [ADV ADJ] cases can be involved in 'normal' comparatives. Therefore the failure of a synthetic form for *amazingly smart* can be attributed to the intervention of the adverb, in line with what the adjacency-based analysis predicts.

3.4.2 Bracketing

The discussion immediately above shows that (1) *more amazingly smart* etc. can be true comparatives, and (2) that the failure of synthetic adjectives in this case follows from the adjacency-based treatment. Kiparsky (2005) attempts to defuse this type of argument by claiming that in e.g. *most amazingly smart*, the only possible bracketing is that in (73a):

(73) most amazingly smart
a. [most [amazingly smart]]
b. [[most amazingly] smart]

According to the first bracketing, what is being picked out is the highest degree of amazing smartness. According to the second, it is the highest degree to which it is amazing that someone is smart that is at issue. The (b)-reading may be less than obvious in this particular example. But the issue is important, because it calls into question the kind of locality or intervention effect in [ADV ADJ] cases, and this is important for a structure/adjacency approach. The background assumption for Kiparsky's proposed reanalysis is that if C/S-formation shows any sort of structural locality properties—i.e. if Deg attaches to anything other than a 'word'—then this is problematic for certain Lexicalist assumptions. If *smarter* exists as a word in the lexicon, it is hard to keep it out of examples like (73).[34] The attempt at rebracketing is apparently supposed to show that Deg never has a syntactic distribution in the case of adjectives that could take synthetic forms, although this point is not spelled out by Kiparsky.

The bracketing-based reanalysis is problematic because in cases like (74), the motivation for saying that [[Deg ADV] ADJ] is the only available bracketing is quite limited:

(74) All my students are amazingly smart, but Mary is the most amazingly smart of them all.

[34] Relatedly one could ask about the comparatives of phrases, like *prouder of his son* from above; Kiparsky does not take these into consideration. See Embick and Marantz (2006).

The most natural interpretation of (74) is not one in which Deg and ADV are bracketed together, but instead one in which it is the degree to which Mary is *amazingly smart* that is functioning as the gradable element: [most [amazingly smart]]. The use of substitution with *so* confirms that this bracketing is in fact available:

(75) All my students are [amazingly smart], but in the end I would say that Mary is the most *so*. =<amazingly smart>

For this argument to go through, it is necessary to show that the anaphoric *so* can be resolved to *amazingly smart* and not just *smart*. With *amazingly smart*, this is complicated by the fact that the [most [amazingly smart]] person is also the smartest person. In order to be absolutely clear, what is needed is an example in which resolving the substitution to [ADV ADJ] and resolving it to [ADJ] mean different things. The example (76) has this property:

(76) All of the inmates in this asylum are dangerously insane, but John is the most so.

In this case, the interpretation resolves *so* to *dangerously insane*, not just *insane*; the person who is the most insane is not necessarily the person who is the most dangerously insane. Accordingly, the substitution in (76) involves <dangerously insane>, confirming that this is (or can be) a contituent.

So the bracketing [most [amazingly smart]] is available in some cases. At the same time, Kiparsky is correct that there is another reading for the string *most amazingly smart* in this example, although it might be difficult without context. An example is presented in (77), where (77a) shows *so*-substitution as an attempt to facilitate this bracketing, and (77b) a paraphrase:

(77) Mary is the [most amazingly] smart student.
  a. All the students in my class are smart, but Mary is the [most amazingly] so. [=smart]
  b. = The degree to which it is amazing that Mary is smart is greater than the degree to which it is amazing that all other students are smart (e.g. because she started school at a late age).

This reading is preferred in other examples (e.g. *more obviously inappropriate* or *most recently arrived*). But this is irrelevant to the argument advanced above. The fact that the bracketing in (77) is possible does not imply that the bracketing [most [amazingly smart]] is impossible. The facts above show that both are possible, and that structures like [Deg [ADV ADJ]] are generated by the syntax.[35]

[35] There are some further questions about interpretation. In some cases, both bracketings are possible, with the alternative intepretations being slightly clearer; the example in (i) has this property:

(i) John was more rudely late than Bill.

According to one reading, John and Bill were both late in a way that was rude, but John arrived even later than Bill. On the second reading, it is possible that they arrived at the same time, and what differs is the extent to which each of their arrivals was rude. The assertion of the second reading is that John arrived in an even ruder fashion; e.g., in addition to coming late, he spilled coffee, etc. These readings correspond to the distinct bracketings identified above:

(ii)a. Reading1: [more [rudely late]] = arrived even later
  b. Reading2: [[more rudely] late] = arrived at the same time, but more rudely

In some other cases it is hard to associate distinct interpretations with the distinct bracketings:

It therefore seems that both bracketings are in principle available. Crucially, the bracketing [Deg [ADV A]] is possible, and, in such cases, a synthetic form cannot be derived.

3.4.3 Left/right

Putting things together, the pattern is as follows. When there is material inside the AP that Deg applies to, for C/S-formation it matters crucially how that material is linearized. For post-head material, even if it is a sister of the adjective, synthetic comparative forms are possible; recall examples like *John is likelier to win the race than Bill is* from (23) above.

On the other hand, material that is linearized to the left of the adjective, such as the adverbs considered above, prevent the creation of synthetic forms. The Deg head and the adjective are never adjacent, hence the rule of Local Dislocation cannot apply.

The left/right asymmetry in C/S-formation follows from the adjacency-based approach in the way that I have demonstrated above.

3.5 Synopsis

The analysis presented in this section builds on the idea that there is significant motivation for a uniform syntactico-semantic treatment of Deg. With this syntactic treatment comes the question of how synthetic C/S-forms are derived. A treatment in terms of Local Dislocation under string-adjacency accounts for the different effects that have been examined above.

This analysis of the analytic/synthetic alternation in C/S-formation takes a specific position with respect to blocking effects. According to the analysis above, there is no sense in which *smarter* blocks *more smart*, or in which *more intelligent* blocks *intelligenter*. Rather, the syntax creates particular structures which may or may not be subject to the rule of Local Dislocation whose properties are detailed above. If the rule applies, a synthetic form is created; if it does not, an analytic form is created. The pattern does not result from competition among possible outputs; rather, it results from the structure required for the syntax and semantics of comparatives and superlatives, along with a rule of the PF component. A generalized syntactic account along these lines holds that the range of word/phrase alternations is delimited by the theory of locality, in particular the theory of the conditions under which head adjunction ("affixation") can take place.

Footnote 35 continued

(iii) Comus is more brutally honest than Reginald.
  a. [more [brutally honest]]
  b. [[more brutally] honest]

The first bracketing asserts that Comus possesses more brutal honesty than Reginald; the second, that his honesty is of a more brutal variety. The differences between these two are rather subtle, at best.

Other diagnostics correlate with the bracketing, and confirm the general point above that there are two different structures in principle possible for *more ADV ADJ* strings. For instance, the presence of a pause associated with prosodic structure correlates with the bracketings discussed above:

(iv)a. more || rudely late
  b. more rudely || late

A further question concerns how this syntactic treatment without word/word or word/phrase competition compares with alternatives that employ blocking or "Poser Blocking" to account for such patterns. While a full discussion of this analysis with respect to alternatives is not for this paper, some aspects of C/S-formation identified above allow for us to see what the outline of such a comparison would look like. The general set of locality conditions implicated in C/S-formation, and in particular the type of "left/right" asymmetry seen in C/S-formation, are accounted for straightforwardly in the approach presented above. This type of asymmetry is not predicted in approaches that implement "Poser Blocking"—blocking of phrases by words—as Hankamer and Mikkelsen (2005) discuss very clearly with reference to a different case study (Bresnan, 1999 also notes problems with a Poser-Blocking treatment of C/S-formation). It looks as if the syntactic/derivational approach makes correct predictions that could be generalized, while there are prima facie problems for blocking-based theories. If a general argument could be made that the syntactic approach makes correct predictions, while competition-based approaches do not, this would be a significant result. See Embick and Marantz (2006) for extensive discussion along these lines.

4 Conclusions

Blocking effects are of interest because their analysis requires explicit architectural claims about the interaction of different aspects of grammatical competence. The nature of these interactions has a particular status within the theory of Distributed Morphology, where, for architectural reasons, blocking in the typical sense cannot be implemented, whether for word/word or word/phrase interactions. This paper's discussion of blocking effects and how they are treated in this theory is centered on two points. The first is that the locus of competition is the *morpheme*: Vocabulary Items compete to determine allomorphy at individual nodes. Other types of competition, between larger objects, cannot be formulated. This leads to the second point, to which most of this paper is devoted. Interactions between words and phrases, of the type found in C/S-formation and discussed above, are the result of combinatory processes in the grammar, in particular the processes that put heads together. This paper was centered on a particular case study. The analysis in Sect. 3 demonstrates what it means for C/S-formation to take place in a uniform syntactic structure. The further question that then arises, concerning the combinatorics, is ultimately one of locality. The conclusion argued for above is that the process responsible for creating synthetic C/S forms is Local Dislocation, an operation that applies under linear adjacency.

A general question for the discussion of blocking effects in analytic/synthetic alternations is whether all such cases involve either linear adjacency or the configuration in which head movement (or local head-to-head Lowering) is possible. A syntactic theory like Distributed Morphology predicts that the cases in which alternations between words and phrases could be found are precisely those delimited by syntactic (including post-syntactic) theories of locality. These points are generalized and compared with alternatives in Embick and Marantz (2006).

As a general comment concerning the types of phenomena treated here, it was shown in the discussion of comparative and superlative formation that understanding a set of ostensibly 'morphological' facts concerning the formation of synthetic C/S-forms requires detailed treatments of the syntax and semantics of comparative

constructions, a rich area of research. I take the analysis above to have demonstrated that it is not possible to talk about putatively morphophonological effects without strong assumptions about other parts of the grammar; morphology is not "by itself". Most of the discussion above involves exploring Non-Lexicalist assumptions for analytic/synthetic alternations, a phenomenon whose treatment requires assumptions in many different subparts of the grammar. If the conclusions advanced above are correct, then progress in understanding such phenomena requires attention to details that are unavoidably syntactic: the analysis of blocking effects must be stated in terms of an approach that takes into account syntactic structures and their interfaces with sound and meaning.

References

Aronoff, M. (1976). *Word formation in generative grammar*. Cambridge, MA: MIT Press.
Bhatt, R., & Pancheva, R. (2004). Late merger of degree clauses. *Linguistic Inquiry, 35*(1), 1–45.
Bobaljik, J. (1994). What does adjacency do? In H. Harley & C. Phillips (Eds.), *The morphology-syntax connection* (pp. 1–32). Cambridge, MA: MITWPL 21.
Bresnan, J. (1973). The syntax of the comparative clause construction in English. *Linguistic Inquiry, 4*, 275–343.
Bresnan, J. (1999). Explaining morphosyntactic competition. In M. Baltin & C. Collins (Eds.), *Handbook of contemporary syntactic theory* (pp. 1–44). Oxford: Blackwell.
Chomsky, N. (1995). *The minimalist program*. Cambridge, MA: MIT Press.
di Sciullo, A. M., & Williams, E. (1987). *On the definition of word*. Cambridge, MA: MIT Press.
Embick, D. (2000). Features, syntax and categories in the Latin perfect. *Linguistic Inquiry, 31*(2), 185–230.
Embick, D. (2003). Locality, listedness, and morphological information. *Studia Linguistica, 57*(3), 143–169.
Embick, D. (2005). Linearization and local dislocation: Derivational mechanics and interactions. to appear in *Linguistic Analysis*.
Embick, D., & Halle, M. (2005). On the status of *stems* in morphological theory. In T. Geerts & H. Jacobs (Eds.), *Proceedings of going romance 2003* (pp. 59–88) Amsterdam: John Benjamins.
Embick, D., & Halle, M. (forthcoming) *Distributed morphology*. Berlin: Mouton de Gruyter.
Embick, D., & Marantz, A. (2006). Architecture and blocking. ms., University of Pennsylvania and MIT.
Embick, D., & Noyer, R. (2001). Movement operations after syntax. *Linguistic Inquiry, 32*(4), 555–595.
Fox, D., & Pesetsky, D. (2005). Cyclic linearization of syntactic structure. *Theoretical Linguistics, 31*, 1–45.
Giegerich, H. (2001). Synonymy blocking and the elsewhere condition: Lexical morphology and the speaker. *Transactions of the Philological Society, 99*(1), 65–98.
Halle, M. (1997). Distributed morphology: Impoverishment and fission. *MIT Working Papers in Linguistics, 30*, 425–449.
Halle, M., & Marantz, A. (1993). Distributed morphology and the pieces of inflection. In K. Hale & S. J. Keyser (Eds.), *The view from building 20: Essays in linguistics in honor of Sylvain Bromberger* (pp. 111–176). Cambridge, MA: MIT Press.
Hankamer, J., & Mikkelsen, L. (2005). When movement must be blocked: A response to Embick and Noyer. *Linguistic Inquiry, 36*(1), 85–125.
Heim, I. (2000). Degree operators and scope. In B. Jackson & T. Matthews (Eds.), *Proceedings of SALT X*. (pp. 40–64). Cornell University, Ithaca: CLC Publications.
Kennedy, C. (1997). Comparison and polar opposition. In A. Lawson (Ed.), *Proceedings of SALT 7* (pp. 240–257). Ithaca: CLC Publications.
Kennedy, C. (1999). *Projecting the adjective: The syntax and semantics of gradability and comparison*. New York: Garland.
Kennedy, C. (2001). Polar opposition and the ontology of "degrees". *Linguistics and Philosophy, 24*, 33–70.
Kennedy, C. (2002). Comparative deletion and optimality in syntax. *Natural Language and Linguistic Theory, 20*(3), 553–621.

Kiparsky, P. (2005). Blocking and periphrasis in inflectional paradigms. *Yearbook of Morphology, 2004*, 113–135.
Kroch, A. (1994). Morphosyntactic variation. In K. Beals (Ed.), *Papers from the 30th regional meeting of the Chicago Linguistics Society: Parasession on variation and linguistic theory* (pp. 180–201). Chicago: Chicago Linguistic Society.
Lechner, W. (1999). Comparatives and DP-Structure. Doctoral Dissertation, University of Massachusetts, Amherst.
Lechner, W. (2004). *Ellipsis in comparatives*. Berlin: de Gruyter.
Marantz, A. (1984). *On the nature of grammatical relations*. Cambridge, MA: MIT Press.
Marantz, A. (1988). Clitics, morphological merger, and the mapping to phonological structure. In M. Hammond & M. Noonan (Eds.), *Theoretical morphology* (pp. 253–270) San Diego: Academic Press.
Marantz, A. (1995). A late note on late insertion. In *Explorations in generative grammar* (pp. 357–368). Seoul: Hankuk Publishing.
Marantz, A. (2001). Words and things. Handout, MIT.
Marantz, A. (2003). Blocking. Handout, MIT.
McCawley, J. D. (1988). *The syntactic phenomena of English*. Chicago: University of Chicago Press.
Newell, H. (2004). A late adjunction solution to bracketing paradoxes. Paper presented at NELS 35.
Pesetsky, D. (1979). Russian morphology and lexical theory. ms., MIT.
Pesetsky, D. (1985). Morphology and logical form. *Linguistic Inquiry, 16*, 193–246.
Poser, W. J. (1992). Blocking of phrasal constructions by lexical items. In I. Sag & A. Szabolsci (Eds.), *Lexical matters* (pp. 111–130). Stanford: CSLI.
Richards, N. (2002). A distinctness condition on linearization. ms., MIT.
Speyer, A. (2005). Bracketing paradoxes as cases of word-internal adjunction. ms., University of Pennsylvania.
Sproat, R. (1985). On Deriving the lexicon. Doctoral Dissertation, MIT.
Williams, E. (2004). Dumping lexicalism. to appear In G. Ramchand & C. Reiss (Eds.), *Oxford Handbook of Linguistic Interfaces*.