

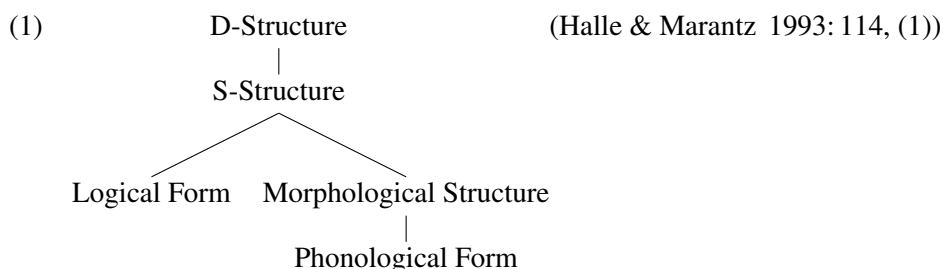
On the Motivation for Phonology-Free Syntax and Multiple Levels in a Derivational Grammar

Benjamin Bruening (University of Delaware)

rough draft, February 18, 2019; comments welcome

1 Introduction

Most current theories of grammar accept Zwicky’s (1969) Principle of Phonology-Free Syntax and strictly separate the information contained in the module of phonology from the module of syntax. Some theories also adopt a serial view of the grammar, with distinct levels and a derivation that proceeds from one to the next. For instance, one currently popular model, Distributed Morphology (Halle & Marantz 1993), proposes the architecture shown in (1). The derivation of a sentence starts with a deep structure (D-Structure) that undergoes various syntactic processes to produce a surface structure (S-Structure). At this point the derivation branches, with one side going to the level of logical form (LF), the input to semantic interpretation, and the other going to phonological form (PF):¹



(This model is simply a modification of the Government and Binding model of syntax proposed in Chomsky 1981.)

Importantly, in Distributed Morphology syntactic derivations manipulate only abstract syntactic features at the levels of D-Structure, S-Structure, and LF. Phonological forms are not inserted until the level of Morphological Structure on the PF branch. Thus, the three purely syntactic levels have no access at all to phonological information. Only at Morphological Structure is such information available. At the level of Morphological Structure, actual morphemes (“Vocabulary Items,” in Distributed Morphology parlance) are inserted, and various manipulations of the syntactic structure can take place. Some of these manipulations can be sensitive to the phonology of the items that have been inserted. This type of serial model achieves strict separation of phonology from syntax by completely removing phonological material from the levels of syntax.

¹In Distributed Morphology, the LF component is strictly separate from Morphological Structure, where Vocabulary Insertion takes place. However, actual semantic interpretation obviously depends on particular Vocabulary Items. This necessitates a later stage where semantic interpretation reunites the output of LF with the output of Morphological Structure. This later stage is rarely discussed and is never represented in diagrams of the Distributed Morphology model, but it is actually quite important (see Harley 2014: 230). This makes the Distributed Morphology model even more complicated than it appears to be in (1).

In this squib, I critically examine the logic behind phonology-free syntax and the strict separation of phonology from syntax. I show that this logic is flawed, and there is no reason to strictly separate phonology from syntax, or to posit multiple levels in a serial model. Instead, we can have a much simpler model of grammar, with all information available from the beginning and syntax and phonology running in parallel. (I will limit my discussion here to derivational models of grammar; non-derivational ones are obviously compatible with parallel syntactic and phonological constraints.)

2 The Logic of Separation

As an example of the way syntax is supposed to be phonology-free, consider head movement of the kind that raises verbs to T(ense) in French or to C(omplementizer) in verb-second languages like German. This head movement process is never sensitive to phonology. For example, no language has a head movement rule that raises only verbs that start with [b] to T. Similarly, no verb-second language raises only verbs with two syllables to C but not verbs with only one syllable. As another example, this time from phrasal movement, no language fronts wh-phrases that end with a sonorant but leaves in situ wh-phrases that end in an obstruent.

Because of this apparent insensitivity of the syntax to phonology, many current approaches propose that phonological properties of lexical items are simply not visible to the syntax. In Distributed Morphology, the syntax deals only with bundles of abstract morphosyntactic features, and actual lexical items are not inserted until the post-syntactic component of Morphological Structure, as explained above. This makes the phonological properties of those items unavailable to the syntax. Similar architectural assumptions of modularity are made in other approaches, as well.

Distributed Morphology relegates more than just phonological information to the post-syntactic component. Following the same logic, Distributed Morphology also locates agreement at the post-syntactic level of Morphological Structure: just as no V-movement rule is sensitive to phonology, there is also no V-movement rule that targets only verbs that agree with a first person subject (cf. Harris 2017: 206–207). Similarly, Halle & Marantz (1993) say that verb class morphemes (e.g., theme vowels in Romance) must also be inserted post-syntactically, because again no V-to-T rule in Romance languages targets only verbs of one class. “By placing them in this part of the grammar, we account for their lack of effect in the syntax or at LF,” say Halle & Marantz (1993: 135).

The logic of this argument is then the following: The grammar includes many different types of information. If certain types of information are never used by some component of grammar like the syntax, then that component must not have access to it. Note that this conclusion rests on the assumption that, if a component does have access to some information, it would necessarily make reference to it. If this assumption is not justified, then the conclusion is no longer valid. This is what I will show here: the assumption is *not* justified. I will also show that there is an alternative explanation for the fact that most syntactic processes are insensitive to phonological information, without the need to separate them.

3 What is Syntax Sensitive to?

In order to see whether the syntax necessarily makes reference to all of the information that is available to it, let us examine what information syntactic operations like head movement should in principle have access to, and what information they actually make reference to. I will start with the second issue: what information operations like head movement *do* make reference to.

For head movement, the answer seems to be only the distinction between phrases and heads, and syntactic category. Head movement ignores phrases and only targets heads, so that distinction is something it must be able to see. Head movement is also typically limited by syntactic category: V-to-T and T-to-C typically

involve only verbs. In English, only auxiliary verbs can undergo either process, and it is reasonable to think that the difference between main and auxiliary verbs is one of (sub)category. Head movement is therefore sensitive to head status and syntactic category. Is head movement sensitive to anything else? The answer seems to be no.

Now the question is, what information should head movement have access to that it is *not* using? The answer is, lots. The difference between a transitive and an intransitive verb must be present in the syntax, because it makes a difference to what phrase structure is projected, what auxiliary is selected (in some languages), what form a causative will take (in many languages), etc. At the same time, however, no language (to my knowledge) has a V-to-T movement rule that targets only transitive verbs and not intransitive verbs. This means that there is information in the syntactic component that a syntactic operation nevertheless ignores, directly contradicting the assumption behind the logic of separation.

Another type of information that the syntax should have access to but head movement ignores is tense on the highest verb. Most Germanic languages make at least a past/non-past distinction on the highest finite verb, but in all of them this distinction is irrelevant to head movement. In English, for instance, subject-auxiliary inversion completely ignores this information, yielding both *Is it raining?* and *Was it raining?* Again, information that should be visible to the syntax is simply irrelevant to a syntactic process.

Let us take as another example phrasal movement: consider raising to subject and raising to object. Raising to subject and raising to object must be sensitive to syntactic category, since they typically target only NPs. They also seem to be sensitive to something like grammatical function or case, since they typically only target certain types of arguments and not others (e.g., obliques). It does not appear that there is anything else that they are sensitive to.

What else could raising to subject and raising to object be sensitive to, in principle? The list is endless: Raising could be limited to targeting only NPs that have a postnominal PP or CP; it could be limited to targeting coordinated NPs; it could be limited to targeting an NP that has a modifying AP within it. Since the syntax is what puts postnominal PPs/CPs and modifying APs together with Ns, and the syntax is what builds coordinated NPs, this information *must* be visible to the syntax; yet few if any syntactic processes ever refer to such things. In fact that is the reason for calling all NPs “NPs,” regardless of their internal makeup: they all behave alike in external distribution. Again, this contradicts the assumption of the logic of separation, according to which operations of a given component would necessarily refer to all information that is available in that component.

As another example of information that is available to the syntax but ignored, Halle & Marantz (1993) conclude that features that distinguish count from mass nouns must be part of the syntax, because they make a difference to determiner selection, quantifier selection, etc. Again, however, no rule of wh-movement, topicalization, or raising targets only count nouns and not mass nouns.

What all of these examples show is that particular syntactic operations *never* make use of the full range of information that must be available to them. In fact, individual syntactic operations seem to be severely limited in what they can make reference to. Both of our examples so far have only made reference to syntactic category and one other piece of information: something like case/grammatical function for raising, and head versus phrase for head movement. Various other phrasal movement processes only seem to care about one additional factor each, in addition to category: topicalization only targets some kind of informational-structural notion borne by a phrase, but does not care about the composition of that phrase; wh-movement only targets wh-phrases but does not care about their internal makeup; relativization is similar. The operation of Merge, combining a selector with its selected argument, or a modifier with its modifiee, only seems to care about category, and particular selectors may pick out some feature in addition, like plural or animate, but Merge is *never* sensitive to the full range of syntactic information that it could be. For instance, Merge never cares whether the merged phrase includes coordination, or a modifier, or anything else that the syntax put together. But again, since the syntax put these phrases together, this is information that the syntax must have access to.

Let me state this as a generalization about syntactic operations:

(2) Generalization Regarding the Sensitivity of Syntactic Operations:

A syntactic operation is sensitive to at most two syntactic features and ignores everything else. One of these features is typically (always?) syntactic category.

As this generalization states, there is a wealth of syntactic information that syntactic operations simply ignore.

I should note that this limitation has always been implicit in the way syntacticians have formulated rules or operations. As just noted, all NPs are grouped together as NPs because they pattern the same, regardless of their syntactic makeup. The same goes for all the syntactic categories: PPs, VPs, CPs, APs, and so on. Considering them to be the same thing, and formulating rules like selection that refer only to these categories, implies that their internal makeup, though visible, is simply not relevant. Prominent in early discussions of subject-auxiliary inversion in English was the recognition that it ignores everything except the structurally highest auxiliary verb, inverting it with the subject NP. The internal makeup of the subject NP, and the content of the highest auxiliary, was explicitly recognized as being irrelevant (this is what is behind the rule being hierarchical rather than linear; Chomsky 1968: 51–52, Chomsky 1971, 26–28; see Berwick et al. 2011). It is redundant of researchers to insist that they need to strictly exclude phonology from the syntax in architectural terms, when in their actual formulations of syntactic processes they have already accomplished that.

The severe limitations on syntactic operations have also been recognized explicitly before, in Relativized Minimality (Rizzi 1990). The point of Relativized Minimality is that particular syntactic operations only care about *one particular thing*: A-bar movement can only see A-bar positions, A-movement can only see A-positions, and head movement can only see head positions. None of them can see other types of positions, nor can they see the actual content of the positions they are concerned with. Importantly, the content these syntactic processes cannot see includes purely *syntactic* content. What Relativized Minimality recognizes is that syntactic operations are *selectively blind*. There is a very limited amount of information that they are sensitive to, and they are not sensitive to anything else.

In the logic of separation, this would have to be taken to motivate even more levels and even more separation. If head movement can see only head status and syntactic category, there must be a module where head movement takes place that has *nothing* but those two pieces of information. There must be another, completely separate, module with nothing but syntactic category and case/grammatical function present, for the mechanism of A-movement. Clearly, this is an absurd model of syntax.

4 Phonology Can be Visible to the Syntax

Once we recognize that particular syntactic operations are radically restricted and selectively blind, do we still need to strictly separate phonology from syntax? The answer is no. If head movement is restricted such that it can only see head status and syntactic category, then it will obviously ignore phonology even if that information is present and visible, since phonology is not head status and syntactic category. Head movement also ignores the difference between transitive and intransitive verbs—completely syntactic information—for the same reason. Similarly, neither agreement features nor conjugation class features need to be shunted off to a different component of grammar. The restriction on head movement to seeing only head status and syntactic category will also block it from seeing agreement features and conjugation class. Similarly, whatever restricts A-movement to seeing only syntactic category and grammatical function/case will block it from seeing phonological information, count versus mass status, declension class, and numerous other pieces of information. In other words, phonological information is not different from syntactic information: the fact

is that syntactic operations ignore almost everything. There is no reason to separate phonology from the bulk of syntactic information.

Thus, it is not necessary to have a model of grammar where phonological information is completely absent from the syntax in order to account for the insensitivity of syntactic operations to phonological information. There is a reason that syntactic operations will ignore it even if that information is available in principle.

5 What Should a Derivational Model Look Like?

Supposing we want a derivational model of the grammar rather than a purely representational one, the multi-level Distributed Morphology model in (1) is clearly far more complex than it needs to be. What should we replace it with? I suggest that we start with something like what Chomsky (1993) proposes: the syntax starts a derivation by taking items from the lexicon and putting them together. These items include with them all the information that they possess: phonological, syntactic, semantic, pragmatic. The derivation proceeds by merging items, building larger and larger items, running various syntactic operations, and ultimately finishing when all operations are complete. The phonology can operate on this same derivation in parallel. There are no levels, only a beginning point and an ending point. The derivation puts all complex items together, whether those complex items are “words” or units larger or smaller than words; as Bruening (2018) argues, there is no need for separate modules of the grammar for word formation and for phrasal syntax.

Representing this model visually, we can replace the numerous levels and branches of (1) with the radically simple architecture shown below:

- (3) The Minimal Derivational Model of Grammar:

start—morphosyntax—end

As described, the derivation proceeds by taking items from the lexicon and putting them together. The lexicon includes both a lexical vocabulary and a functional one, and a well-formed derivation will have to include a certain set of functional items.

There are of course going to be numerous complexities, and many questions of implementation will arise, but I suggest that this minimal derivational model is all we need. All information is present throughout, with different components of the grammar—syntax, semantics, pragmatics, phonology—operating in parallel. Future work in derivational approaches like Distributed Morphology should attempt to flesh out this minimal derivational model, and should give up the model in (1).

6 Could Syntactic Processes be Sensitive to Phonology?

The generalization arrived at above was that syntactic operations pay attention to at most two features, one of which is (usually/always) syntactic category. This leaves open the possibility that the second one might be phonological in nature. Indeed, it may well be the case that such processes exist. Consider heavy shift. One possible way of looking at heavy shift is that it is a syntactic operation that cares about the two features of syntactic category and phonological weight. This is a possibility that theories with strict separation would never allow, but the current approach does. There are other phenomena that might potentially also be instances of syntactic processes that are sensitive to phonology, for instance the phonological agreement described by Sande (2016). I suggest that a fruitful line of inquiry will be giving up the strict separation of syntax and phonology, and pursuing analyses of these phenomena where the syntax has direct access to phonology.

7 An Argument that Syntax MUST See Phonology

As noted above, Bruening (2018) argues strongly that there is no separate component of the grammar for word formation. All complex objects are put together by the same system, the syntax. This is a position that has been argued for before, and it is an assumption of Distributed Morphology, as well. Assuming that it is correct, there is then a very simple argument that phonological information *must* be available to the syntactic component. This argument involves alternations like that found with the English comparative and superlative.

The English comparative and superlative are either morphological or periphrastic, depending on a variety of factors. One factor is phonological: only adjectives that are a single prosodic foot can form *-er* comparatives and *-est* superlatives (*smarter, smartest; *intelligenter, *intelligentest*). Others form a periphrastic comparative/superlative with *more/most* (*more intelligent, most intelligent*).

The argument is very simple: If it is the syntax that puts together *smart* and *-er/-est*, then the syntax has to be able to see the phonology, otherwise it would blindly attach *-er/-est* to any adjective, regardless of its phonological weight.

This conclusion is unavoidable. Working within Distributed Morphology, Bobaljik (2012) was forced to posit a diacritic in the syntax that has the effect of producing synthetic comparatives with short adjectives (*smarter*) but analytic comparatives with long adjectives (*more intelligent*). This diacritic is just a stand-in for phonological properties. In other words, phonological information *must* be part of the syntax; in order to avoid having phonological information in the syntax in the already overly complex model of Distributed Morphology, the additional complexity of a mediating feature has to be introduced. Since the mediating feature just points to phonological information, it is clear that what we actually need is the phonological information.

Thus, if the syntax is what builds morphological objects like *smarter*, it simply *must* be able to see phonological information.

8 Further Advantages to Rejecting Separation

A model with all information present and together has numerous advantages, including simplicity. As just noted, Bobaljik (2012) was forced to posit a diacritic in the syntax for comparatives and superlatives. This is totally unnecessary in the simpler model advocated here. If all phonological information is present from the beginning, it is available to the (morpho-)syntax and can be accessed directly in forming comparatives and superlatives.

Distributed Morphology also has to divide actual lexical items into two distinct lists: the list of phonological forms (Vocabulary Items), and the list of “encyclopedic” information, or semantic information. This distinction can be dispensed with, if all lexical items have all phonological, semantic, pragmatic, and other information present with them throughout the derivation.

The architectural assumptions of Distributed Morphology have also led to numerous suboptimal analyses and even to clearly incorrect claims. Marantz (1997) concluded from the Distributed Morphology architecture that there could be no suppletion for lexical roots, a conclusion that is clearly false (e.g., Harley 2014). Reasoning from the same assumptions, Harley (2014) concluded that terminal nodes for lexical roots in the syntax have no content except for a numerical pointer, which points to the particular lexical root on the list of Vocabulary Items and the Encyclopedia. This is equivalent to having the lexical root present from the beginning, except that it adds an additional theoretical device, the numerical pointer. This is also completely unnecessary in a simpler model where all information is present and together.

9 Conclusion

In this squib, I have shown that the assumption underlying the argument for the strict separation of phonology and syntax is invalid. We do not need to strictly separate phonological information from the syntax in order to account for the fact that most syntactic processes make no reference to phonological information. The fact is that most syntactic processes make no reference to most syntactic information, too. Syntactic processes are severely constrained, in such a way that their inability to refer to phonology is just a sub-case of all the other information they are also forced to ignore. Conversely, if there is only a single component of grammar for both the morphology and the syntax, then that component *must* have access to the phonology. I have argued that the simplest possible derivational model is the one that is best supported: one where a single generative engine takes elements from the lexicon and puts them together, with all phonological, morphological, syntactic, semantic, and pragmatic information present and together from the beginning. Phonology and syntax can operate on this derivation in parallel. If we accept this conclusion, then we can have a maximally simple model of the architecture of the grammar, and this will have numerous benefits.

References

- Berwick, Robert C., Paul Pietroski, Beracah Yankama & Noam Chomsky. 2011. Poverty of the stimulus revisited. *Cognitive Science* 35. 1207–1242.
- Bobaljik, Jonathan David. 2012. *Universals in comparative morphology*. Cambridge, MA: MIT Press.
- Bruening, Benjamin. 2018. The lexicalist hypothesis: Both wrong and superfluous. *Language* 94. 1–42.
- Chomsky, Noam. 1968. *Language and mind*. New York: Harcourt, Brace, Jovanovich.
- Chomsky, Noam. 1971. *Problems of knowledge and freedom*. London: Fontana.
- Chomsky, Noam. 1981. *Lectures on government and binding*. Dordrecht: Foris.
- Chomsky, Noam. 1993. A minimalist program for linguistic theory. In Kenneth Hale & Samuel Jay Keyser (eds.), *The view from building 20: Essays in linguistics in honor of Sylvain Bromberger*, 1–52. Cambridge, MA: MIT Press.
- Halle, Morris & Alec Marantz. 1993. Distributed morphology and the pieces of inflection. In Kenneth Hale & Samuel Jay Keyser (eds.), *The view from building 20: Essays in linguistics in honor of Sylvain Bromberger*, 111–176. Cambridge, MA: MIT Press.
- Harley, Heidi. 2014. On the identity of roots. *Theoretical Linguistics* 40. 225–276.
- Harris, Alice C. 2017. *Multiple exponence*. Oxford: Oxford University Press.
- Marantz, Alec. 1997. No escape from syntax: Don't try morphological analysis in the privacy of your own lexicon. In Alexis Dimitriadis, Laura Siegel, Clarissa Surek-Clark & Alexander Williams (eds.), *Proceedings of the 21st annual penn linguistics colloquium* (Penn Working Papers in Linguistics 4:2), 201–225. Philadelphia: University of Pennsylvania.
- Rizzi, Luigi. 1990. *Relativized minimality*. Cambridge, MA: MIT Press.
- Sande, Hannah. 2016. An interface model of phonologically determined agreement. In Kyeong min Kim et al. (ed.), *Proceedings of the 33rd west coast conference on formal linguistics*, 339–350. Somerville, MA: Cascadilla Proceedings Project.

Zwicky, Arnold M. 1969. Phonological constraints in syntactic descriptions. *Papers in Linguistics* 1. 411–463.

Department of Linguistics and Cognitive Science
University of Delaware
Newark, DE 19716
(302) 831-4096
bruening@udel.edu