

More Aspects of Aspects

Howard Lasnik, reply by Anna Maria Di Sciullo

To the editors:

I am pleased to have this opportunity to comment on Anna Maria Di Sciullo's essay as I share her admiration for this classic book by Noam Chomsky. Like her, I regard *Aspects of the Theory of Syntax* as a landmark in the investigation of human language in particular and human cognitive capacity in general. Di Sciullo is spot on in emphasizing what *Aspects* has to say about language acquisition. For my money, the first chapter of that book is still the most compelling statement regarding language acquisition as the core of generative grammar. Di Sciullo has done quite a good job at a very difficult task: explaining to an audience largely of non-specialists what was so important about this book. Executing this task necessarily required considerable simplification. Since I have been studying this book since three years after its 1965 publication,¹ I will attempt to elaborate on some such points and clarify a few, while avoiding, to the extent possible, formidable technicalia.

In giving some earlier Chomskyan background, Di Sciullo writes, "In *Syntactic Structures*, Chomsky identified creativity with the recursive structure of a natural language." I don't find such a discussion, or any discussion at all, of creativity in that earlier groundbreaking book. To the extent that it is there, it is deeply between the lines. To be clear, I don't doubt that Chomsky had something like this in mind at the time; he just didn't express it in that book. Something related, though without explicit mention of recursion, is expressed in Chomsky's *The Logical Structure of Linguistic Theory (LSLT)*, completed in 1955, but not published until 1975. On the first page he writes:

A speaker of a language has observed a certain limited set of utterances in his language: On the basis of this finite linguistic experience he can produce an indefinite number of new utterances which are immediately acceptable to other members of his speech community.²

Discussing the Standard Theory and what it carried over from *Syntactic Structures*, Di Sciullo correctly remarks that phrase structure rules "generate hierarchical structures." I would like to elaborate on this point. The phrase markers do indeed characterize hierarchy—what the constituents

are, and what the labels of those constituents are—as well as NP, VP, etc., and linear order. This latter point is interesting because much of Chomsky's more recent Minimalist work has linear order added to purely hierarchical structure only in the transition to phonology, as Di Sciullo points out in her "Open Questions" section. For syntax itself, hierarchy is all. But in *Aspects*, Chomsky argued that linear order *is* introduced by the phrase structure component, via the concatenation implicit in the rules. He considers and rejects the alternative suggestion that component produces sets: "[T]he evidence presently available is overwhelmingly in favor of concatenation-systems over set-systems, for the theory of the categorial component."³

As Di Sciullo notes, *Aspects*, following *Syntactic Structures*, had another syntactic module: transformational rules, which map phrase markers onto phrase markers. She also notes that transformational rules had been introduced earlier by Zellig Harris, Chomsky's mentor. She suggests that the difference was that "in *Syntactic Structures* they were, for the first time, embedded in a purely formal context." I don't see it quite that way. Harris elaborates on his theory of transformations in works published in 1952 and 1957.⁴ The statements of transformations in these papers seem about as formal as those in *Syntactic Structures*, which cites them both, and are partly an informal sketch of a small portion of the content of *LSLT*. If Di Sciullo actually has in mind *LSLT*, that's a different matter. In *LSLT*, Chomsky presents a complete set theoretic formalization of a grammar of English, going far beyond anything in Harris's papers or in *Syntactic Structures*. I do agree, however, that there was a big difference between Harris's transformations, in any of his work, and those of Chomsky, in any of his work. For the former, transformations were relations between sentences while for the latter, they were, as Di Sciullo says, relations between phrase markers, leading, as she also notes, to the crucial distinction between underlying form—deep structure in *Aspects*—and more superficial form—surface structure. This difference goes along with the divergent goals of Harris and Chomsky. Harris's transformations were devices for normalizing texts. Chomsky's were parts of generative grammars. Chomsky offers some nice discussion of this difference in the preface to the published 1975 version of *LSLT*.

I suppose it is true that “Chomsky electrified the community of linguists by persuasively arguing that the surface structures of a natural language are no good guide to its deep structures.” But the breakthrough behind that argument was the new idea that there actually *is* such a thing as deep structure, that sentences have abstract multiple structures, related by transformations. Unlike Di Sciullo, I don’t really see how the famous sentence *Colorless green ideas sleep furiously* was appealed to by Chomsky in establishing “the distinction between deep and surface structures” since in that example there is no very interesting difference between the two levels of representation. Not much of relevance happens transformationally. Di Sciullo’s other examples, such as *John is easy to please*, are much more revealing in that regard, especially as compared with *John is eager to please*. As Di Sciullo illuminatingly indicates, in the former *John* is the understood object of *please*, which is clearly not true of the latter. This contrast is opaque on the surface, but detectable in deep structure. The deep and surface structures of the *eager* example are quite similar, while those of the *easy* one differ sharply, with the deep structure being much like that of *It is easy to please John*, and the surface structure arising from a transformation that came to be called “tough movement.”

Di Sciullo’s other related example pair is even more intriguing. *John’s eagerness to please is fine*, but not *John’s easiness to please*. A relevant descriptive generalization might be that deep structures can become noun phrases while derived structures cannot. Neither the classic theory of *LSLT* and *Syntactic Structures* nor the Standard Theory of *Aspects* was able to account for this generalization. This was because it was assumed that the relation between a verb and corresponding nominal was transformational, and there was nothing in principle preventing the *easiness* transformation from following tough movement. Five years after *Aspects*, Chomsky did present a solution.⁵ In that 1970 paper, he rejected the former generative view that nominalizations are created by transformations in favor of the idea that they are in the lexicon. Since lexical insertion happens at deep structure, it necessarily precedes all transformations, including tough movement, and thus there can’t be nominalizations of derived structures.

But there was a fly in the ointment. As noted in by Chomsky in 1970, there are constructions that look like nominalizations of some derived structures, such as passive constructions like *Rome’s destruction by the barbarians*. Chomsky proposed that here we have a passive-like transformation applying in a noun phrase, rather than a nominalization of a passivized sentence. But then we need a stipulation for the unacceptable *easiness* example: Passive can apply in a noun phrase, but tough movement cannot.

As far as I know, there is still no really satisfying solution to this problem. On the topic of active versus passive sentences, I question Di Sciullo’s claim that passives cannot “be handled by phrase structure rules, unless the

phrase-structure rules are themselves allowed to increase without limit.” I don’t see why this is any more true of passives than of actives. What *is* true is that phrase-structure rules alone cannot account for the felt relatedness between actives and passives; nor for implicational generalizations like if $NP_1 V NP_2$ is a sentence, then so is $NP_2 be V+ed by NP_1$; nor for the fact that an NP that can be the object of an active can be the subject of the corresponding passive, and one that can’t can’t:

Mary kicked the ball / The ball was kicked by Mary

*Mary kicked sincerity / *Sincerity was kicked by the Mary.

The conclusion, though, is the same as Di Sciullo’s: transformations are well motivated in this framework.

In discussing transformations, Di Sciullo correctly emphasizes their structure-dependent nature. They manipulate units of structure, often called constituents. There are no transformations, for instance, exchanging the third and fifth words. *Aspects* has important discussion of this property of transformations.⁶ One classic case of structure dependence was first discussed in these terms three years after *Aspects*.⁷ Chomsky considers a few conceivable versions of the process relating declaratives with interrogatives. One version gives pairs like

[The subjects who will act as controls] will be paid

Will [the subjects who will act as controls] — be paid?

Here, the first auxiliary verb following the subject, marked with brackets, has moved to the front. But there are never processes such as to move the left-most occurrence of an auxiliary to the front, which would yield the completely impossible

*Will [the subjects who — act as controls] will be paid?

Di Sciullo gives a variant of this argument, but I find Chomsky’s version a bit clearer.

A quick comment about recursion might be in order. Di Sciullo is correct in indicating that in *Aspects* the recursive property of the grammar is relocated from the transformational component—that is, generalized transformations—to the phrase structure component. It isn’t clear to me, however, that “[f]ewer symbols were now required.” Fewer rule types for sure: the previous theory had phrase structure rules, singular transformations operating on single trees, and generalized transformations combining multiple trees into one. The Standard Theory eliminates the third type of operation, which was one of Chomsky’s main arguments in *Aspects* for the change. But I’m not sure how to tell if there are, as a result, fewer symbols. And it seems a bit misleading to claim that “[w]ith recursion, there is in *Aspects*, a return to the creativity of

language” since recursion was always there—just in a different module.

Next, a word on syntax versus semantics. I completely agree that Chomsky clearly distinguished the two, and that in *Syntactic Structures* he used the example *Colorless green ideas sleep furiously* to argue that a sentence can be nonsensical but grammatical. On a more technical facet of the question, with respect to *Aspects*, Di Sciullo remarks that “Chomsky also proposed to distinguish between categorical and semantic selectional features.” The context is the fact that the verb *frighten* “requires a [+animate] object,” a property expressed in the lexical entry of the verb. The question arises whether this requirement is treated in *Aspects* as semantic or syntactic. Chomsky did indeed discuss the issue but didn’t come to a conclusion.⁸ If anything, he hinted that such requirements are *syntactic*. Near the end of that discussion, he wrote,

Thus it seems to me that examples such as (15) [an example about frightening sincerity] do not present a particularly strong argument for removing selectional rules from the syntactic component and assigning their function to the interpretive semantic rules.

In addition to such selectional restrictions determining the sort of NP a verb can take as its object, there were in the lexical entries of verbs in the *Aspects* model subcategorization restrictions, indicating whether a particular verb takes a direct object at all. I assume this is what Di Sciullo is alluding to in her mention of context sensitivity in the distinction between *frighten* and *sleep*, with the former demanding a direct object and the latter normally not tolerating one. “Context-sensitive rules,” she writes, “could well be used to settle the distinctions between *frighten* and *sleep*, but only by adding complexity to the grammar.” But *Aspects* does use context-sensitive rules for this, what Chomsky calls lexical insertion transformations. I might be misunderstanding Di Sciullo’s point here, because a little earlier she said what I just said: “The grammar also contains context-sensitive rules: $A \rightarrow Z / X_Y$, where X or Y are not null. These rules serve to insert lexical items into phrase markers.” The lexicon is the repository of all of these selectional and subcategorization restrictions, along with, of course, phonological underlying representations and meanings, in some format. Thus, it contains many of the idiosyncrasies of particular languages. But I think it goes too far at this stage to say, “Beyond the lexicon, every human language is governed by the same structures of universal grammar, and in this sense, Chomsky argued, there is only one human language.” This became a major theme in the government-binding era of the 1980s and into the Minimalist era of the 1990s and 2000s.⁹ But in the 1960s, Chomsky still assumed that, while there are some universal grammatical principles, the grammatical rules of languages can differ, sometimes significantly. In *Aspects*, Chomsky outlines what a theory of linguistic structure

must contain if it is to account for language acquisition. Such a theory must, among other things, contain “an enumeration of the class G_1, G_2, \dots of possible generative grammars” and “a method for selecting one of the (presumably, infinitely many) hypotheses that are allowed by [that enumeration] and are compatible with the given primary linguistic data.”¹⁰ This latter method is referred to as an evaluation measure. Obviously, if there is only one possible grammar, no evaluation measure would be needed.

On the topic of the lexicon, I like Di Sciullo’s point that categories like *noun*, *verb*, and so on, receive no external definitions in Chomsky’s work. They merely determine syntactic distribution via the phrase structure and transformational rules. One quibble: To the extent that Chomsky used features like [+N], [+V], etc., they were actually not binary features; rather, they were privative ones. Until the mid-1970s, [+N] just meant *noun*, and there was no [-N]. Di Sciullo was correct in citing Roman Jakobson for binary features; syntax was a generation or so behind phonology in this regard. Another quibble: Di Sciullo writes, “In Chomsky’s essay ‘Remarks on Nominalization,’ the binary syntactic features [\pm N] and [\pm V] are used to define the major syntactic categories, N: [+N, -V], V: [-N, +V], ADJ: [+N, +V], P: [-N, -V].” Di Sciullo is in excellent company in giving this citation as probably 99% of writings on this topic give that same citation. But it is completely incorrect. That paper has neither this categorization, nor any such categorization. As I implied just above, the features there are privative, not binary. Further, not only is the standard but mistaken categorization not there, it actually *couldn’t* be there.¹¹ Look at what Chomsky says in *Remarks*:

It is quite possible that the categories noun, verb, adjective are the reflection of a deeper feature structure, each being a combination of features of a more abstract sort. In this way, the various relations among these categories might be expressible. *For the moment, however, this is hardly clear enough even to be a speculation* [emphasis added].¹²

To conclude, I’ll reiterate that Di Sciullo presented an effective overview of what made *Aspects* so important. I suspect that it will motivate some readers to take a look, or a second or third look, at *Aspects* itself, which I am sure they will find rewarding. I hope they will also find useful some of the elaborations and clarifications I have provided.

Howard Lasnik

Anna Maria Di Sciullo replies:

I am very pleased to receive comments from Howard Lasnik, who has produced influential work in generative grammar since its early stages.¹³

It is difficult not to consider *Aspects* as a landmark publication. As Lasnik points out in his letter, it is nonetheless a challenge to expose the most salient insights of this book for a broad audience in a short essay. I agree with his observations and clarifications. In this response, I will draw attention to another dimension of the book that is characteristic of the generative enterprise.¹⁴ In doing so, I will also briefly address some of his comments.

It is important to note that the basic principles of transformational generative grammar were first explained by Noam Chomsky in *The Logical Structure of Linguistic Theory* (*LSLT*, 1955) and brought to a broader audience in his *Syntactic Structures* (1957).¹⁵ As Lasnik points out, this prior history should be acknowledged. By the same token, it would come as no surprise that the notion of recursion, introduced in *LSLT*, featured in the Standard Theory. In *Aspects*, sentence recursion is introduced by the rewriting rules of the base, and not by transformations, as had been suggested previously.¹⁶ This revision led to a simplification of the theory of transformational grammar.

In *Aspects*, Chomsky tentatively accepted Jerrold Katz and Paul Postal's suggestion that syntax and semantics are systematically connected at the level of deep structure.¹⁷ Lasnik is correct in saying that there are no significant syntactic differences between deep and surface structure with examples such as *Colorless green ideas sleep furiously*, since no transformations apply in their derivation, contrary to examples such as *John is easy to please*. The difference in their interpretation is computed at the level of deep structure and not at the level of surface structure.

Empirical and theoretical arguments led to the revision of the grammatical architecture that Chomsky proposed in *Aspects*. In government and binding theory,¹⁸ introduced during the 1980s, semantic interpretation also took place at logical form, derived from surface structure by covert displacement operations, such as quantifier raising.¹⁹ Although this new theory provided important insights on language and a large coverage of language variation, it became increasingly complex with many modules, levels of presentations, and conditions on representations. In the early 1990s, these issues led to the development of the minimalist program: a research space aimed to reduce grammar to its simplest form.²⁰

In *Aspects*, as in the generative enterprise, given observational and descriptive adequacy, working hypotheses proposed in previous stages of the theory are simplified or eliminated. This not only satisfies basic methodological principles in science, but also adds to our understanding of the computational properties of the language internal to the mind.

Howard Lasnik is a Distinguished University Professor in the Department of Linguistics at the University of Maryland.

Anna Maria Di Sciullo is Professor of Linguistics at the University of Quebec at Montreal.



1. Not 1964, as Di Sciullo strangely has it in her first sentence. Another chronological curiosity occurs in her mention of David Pesetsky's striking just the right note "writing almost thirty years later." But Pesetsky's comments were in 2015, so fifty years later.
2. Noam Chomsky, *The Logical Structure of Linguistic Theory* (New York: Springer, 1975).
3. Noam Chomsky, *Aspects of the Theory of Syntax* (Cambridge, MA: MIT Press, 1965), 125.
4. See Zellig Harris, "Discourse Analysis: A Sample Text," *Language* 28, no. 4 (1952): 474–94, doi:10.2307/409683; Zellig Harris, "Co-Occurrence and Transformation in Linguistic Structure," *Language* 33, no. 3 (1957): 283–40, doi:10.2307/411155.
5. Noam Chomsky, "Remarks on Nominalization," in *Readings in English Transformational Grammar*, eds. Roderick Jacobs and Peter Rosenbaum (Waltham, MA: Ginn, 1970), 184–221.
6. Chomsky, *Aspects*, 55–56.
7. Noam Chomsky, *Language and Mind* (New York: Harcourt Brace Jovanovich, 1968), 61–62.
8. Chomsky, *Aspects*, 153–59.
9. The government-binding era was initiated by Chomsky. See Noam Chomsky, *Lectures on Government and Binding* (Dordrecht: Foris, 1981). For Chomsky's work on minimalism, see Noam Chomsky, *The Minimalist Program* (Cambridge, MA: MIT Press, 1995).
10. Chomsky, *Aspects*, 30–31.
11. So where did Chomsky introduce this binary feature-based categorization? As far as I can tell, the first Chomsky publication containing it was published in 1977. Three years earlier, it appeared in an unofficial publication: a transcript of a series of lectures presented at the Linguistic Society of America summer school at University of Massachusetts, Amherst. See Noam Chomsky and Howard Lasnik, "Filters and Control," *Linguistic Inquiry* 8, no. 3 (1977): 433; and Noam Chomsky, *The Amherst Lectures* (Paris: Université de Paris VII, 1974).
12. Noam Chomsky, "Remarks on Nominalization."
13. E.g., Chomsky and Lasnik, "Filters and Control."
14. Peter Graff and Coppe van Urk, eds., *Chomsky's Linguistics: The Goals of the Generative Enterprise* (Cambridge, MA: MIT Press, 2012).
15. Chomsky, *The Logical Structure of Linguistic Theory*; and Noam Chomsky, *Syntactic Structures* (The Hague: Mouton, 1957).
16. Chomsky, *Aspects*, 143.
17. Jerrold Katz and Paul Postal, *An Integrated Theory of Linguistic Descriptions* (Cambridge, MA: MIT Press, 1964); and Chomsky, *Aspects*, 141, 145.

18. Noam Chomsky, *Lectures on Government and Binding: The Pisa Lectures* (Dordrecht: Foris, 1981).
19. Robert May, *Logical Form: Its Structure and Derivation* (Cambridge, MA: MIT Press, 1985).
20. Chomsky, *The Minimalist Program*. 60.

DOI: 10.37282/991819.22.6

Published on January 4, 2022

<https://inference-review.com/letter/more-aspects-of-aspects>