

This PDF includes a chapter from the following book:

Born to Parse

How Children Select Their Languages

© 2020 Massachusetts Institute of Technology

License Terms:

Made available under a Creative Commons
Attribution-NonCommercial-NoDerivatives 4.0
International Public License

<https://creativecommons.org/licenses/by-nc-nd/4.0/>

OA Funding Provided By:

The open access edition of this book was made possible by generous funding from Arcadia—a charitable fund of Lisbet Rausing and Peter Baldwin.

The title-level DOI for this work is:

[doi:10.7551/mitpress/12799.001.0001](https://doi.org/10.7551/mitpress/12799.001.0001)

Notes

Chapter 1

1. We return to the distribution of VP ellipsis in §4.2, where we will see that the restrictions on deletion in abstract structures captures a wide range of apparently unrelated phenomena.
2. Footnotes omitted.
3. Footnote omitted. Musso et al.'s results were anticipated in Smith and Tsimplici's 1995 investigations into a linguistic savant's learning of artificial languages, which differed greatly depending on whether the language conformed or not to the demands of UG.
4. The nonoccurrence of **I wonder whether has Kim visited Washington*, where subject inversion causes the structure to crash, is explained by the fact that two words, the "complementizer" *whether* and the preposed *has*, need to be in the same position, namely in the C that heads the clausal phrase, CP, that dominates *Kim visited Washington*. But there is room for only one of them.
5. Jim Higginbotham was the source of this metaphor, at the 1979 workshop in Pisa where Chomsky first presented his new thinking about government and binding (Chomsky 1981a).
6. For discussion, see Chomsky 1995: 212, n. 4.
7. The thinking seems to be that, if one takes biolinguistics seriously, one must use the current primitives of biologists, and not invoke features, *even if features are needed for linguistic description*. This prejudice

recalls debates between Chomsky and Piaget, when Piagetians argued that postulated UG constraints simply cannot be innate if specific to language, regardless of how useful they are analytically, unless they are stated in terms familiar to biologists (Piattelli-Palmarini 1980).

8. It is worth noting that the metric is technically and conceptually flawed insofar as it is based on an assumption that grammars with a greater number of parameters set correctly will be the fittest, the most successful in parsing/generating incoming data. Dresher 1999: 54–58 demonstrates that this assumption is false, that there is no smooth correlation between accuracy in what is parsed/generated and the number of parameters set correctly. See Lightfoot 2006a: 74 for discussion.

9. For good, wide-ranging discussion of parameters, see Karimi & Piattelli-Palmarini 2017. In that volume, some papers argue for dispensing with formal parameters at the level of UG, but Cinque 2017 and Rizzi 2017 offer responses to some of the criticisms. Epstein, Obata, and Seeley 2017 has proposals that overlap with Lightfoot 2017b in advocating an open UG. From the same volume, Longobardi & Guardiano 2017 is motivated by similar concerns to ours and seeks to replace the Principles and Parameters approach with a simplified model of the language faculty, which eliminates parameters altogether from UG, replacing them with a few abstract variation schemata. Also see Baker 2001, a very insightful discussion of parameters analogized to the elements of chemistry.

10. Berwick 1985 keys language acquisition to the products of a parsing device, and Fodor 1998a tackles problems in generating multiple parses. Parsing is central also for our children in their language acquisition, but in contrast and perhaps overambitiously, we postulate no specific parsing procedures distinct from I-language elements.

11. One point worth noting here, to be explored further in chapter 5, is that by not invoking UG-defined parameters we have no way to exclude wild impossibilities: a preposition followed by an IP complement, for example. However, such impossibilities would never be triggered, given that a person's E-language is generated by sets of I-languages in other speakers of the community. If such structures could never be triggered, they do not need to be specifically excluded. People who are contributing to an individual's ambient E-language are using exactly the same mechanisms as people parsing that E-language, namely I-languages subject to

UG constraints and to the triggering effects of their experience; they are not using parsing procedures independent of I-languages.

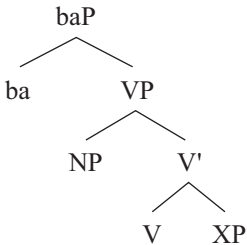
Chapter 2

1. Many people have contributed to our current understanding of these two phase transitions. Roberts 2007 gives a good, detailed textbook account of both changes in I-languages, though viewing them as changes in parameter settings.
2. Lightfoot 2006a: 57–61 discusses the Binding Theory, a vast improvement on earlier analyses of the referential properties of DPs, but a theory that nonetheless raises (solvable) learnability issues, which are not addressed in the literature prior to Lightfoot 2006a. See §4.1 here.
3. For more discussion of how a particular structural element may be triggered by quite different PLD in different I-languages, see Lightfoot 2006a: 123–136.

Chapter 3

1. This is not a terminological issue: thinking in terms of I-languages “changing” or being “restructured” has misled linguists into postulating DIACHRONIC PROCESSES, where one I-language becomes another by some formal operation, and then asking about the nature of those processes, whether they simplify the grammars or make them more efficient or drive them to a different type or something else along those lines. This is a topic for another discourse (Lightfoot 2017a).
2. Ambient E-language typically has several sources, including multiple I-languages, as emphasized by Aboh 2015, 2017.
3. For a more up-to-date and radical approach to ideas about abstract case, see Preminger 2014.
4. Caveat lector: the important thing is to *derive* the change in meaning of psych verbs from a property that would affect language acquisition. Here we derive the change from the loss of inherent case due to the loss of morphological case. This involves general claims about the relationship between morphological and inherent case that may need revision or elaboration. See Preminger 2014.

5. For critical discussion, see Lightfoot 1979: 224.
6. Meanwhile Huang, Li, and Li 2009: §5.4.1, (53) puts forward the very similar but somewhat simpler:



Chapter 4

1. For example, Elbourne investigates the Saliency Hypothesis, the idea that the different behavior of referential and quantificational antecedents does not reflect a real difference in the way that bound and referential pronouns are analyzed with respect to Principle B; rather, it arises because children interpret pronouns as referring to the most prominent characters.
2. Elbourne 2005 and Conroy et al. 2009 dispute Thornton and Wexler's idea that Principle B violations occur only with nonquantificational antecedents. Conroy et al.'s paper constitutes a major clarification of the complex literature on the alleged "delay-of-Principle B effect."
3. Lobeck 1995 and Zagana 1988 offer good Government and Binding accounts of VP ellipsis in terms of the Empty Category Principle of more than thirty years ago; this section adopts the spirit of those analyses in requiring a host to license ellipsed VPs and, in doing so, draws on Lightfoot 2006b.

As a simple illustration of the apparent role of a host for deletion sites, Potsdam 1997 observes the distinction between (ia) and (ib,c), where *do* and *not* appear to license an ellipsed VP; Potsdam notes a similar distinction between (iia) and (iib), where *to* licenses an ellipsed VP.

- (i) a. *It is possible to eat this fruit, and we recommend that you
_{VP}[eat this fruit].

- b. It is possible to eat this fruit, and we recommend that you do $_{VP}$ [eat this fruit].
- c. It is possible to eat this fruit, and we recommend that you not $_{VP}$ [eat this fruit].
- (ii) a. *Kim began singing a song before Jim began $_{VP}$ [sing a song].
- b. Kim began to sing a song before Jim began to $_{VP}$ [sing a song].

4. I assume here that restrictive relative clauses are complements to nouns, distinguishing (10b, 11b); I will take nonrestrictive relatives to be noncomplements, that is, adjuncts. We need a syntactic distinction between restrictive and nonrestrictive relative clauses, and restrictive relatives have some properties of complement clauses. What I am calling complement structures may be captured through Richard Kayne's 1994 raising analysis of restrictive relatives, where a restrictive relative headed by *that* is the complement of D and the head raises out of the relative clause. For a good discussion of relative clauses and the problems they pose for modern theories of phrase structure, see Borsley 1997.

5. Bošković and Lasnik 2003, adopting ideas from Pesetsky 1991 (an unpublished extension of Pesetsky 1995), treats null complementizers as resulting from a phonological affixation operation. For Bošković and Lasnik, affixation requires adjacency, but the data of (11) show that head-complement relations are also crucially involved.

6. Notice that *Which man did Jay introduce to Ray and Jim to Tim?*, analogous to the ill-formed (14c), is well-formed. Here only one *wh*-phrase is overt and it moves across the board. One way of thinking of this is that across-the-board movement takes place on a three-dimensional structure before the two clauses are linearized; at that point *which man* is the complement of *introduce* (Williams 1978).

7. I also adopt the proposals of Nunes 2004, namely that deletion of the copied element follows from the linearization of chains. Linearization is a phonological operation that converts a syntactic structure into a sequence of items in consonance with the Linear-Correspondence Axiom of Kayne 1994. The two *whats* in a structure like (i) are nondistinct, and this leads to ordering contradictions. *What* must precede *buy*, for instance, but *what* must also follow *buy*. That is a contradiction: *x* cannot both precede and follow *y*.

- (i) [what [$_{IP}$ [$_{IP}$ you did buy ~~what~~]]]

It is this failure to yield a linear order that renders the structure ill-formed—unless one of the *whats* is deleted; and it must be the lower *what*, for reasons of the Binding Theory. So, the fact that there can be no chains in the phonology with more than one overtly realized link entails that the lower *what* in (i) must be deleted. Nunes offers a rich analysis, noting exceptional cases where multiple *wh*- items are pronounced; there he shows that it is only intermediate copies that may be pronounced, not the lowest copy, and, indeed, that these pronounced copies have clitic-like qualities (see Nunes 2004: 38–43 for discussion).

8. There are many interesting distinctions at work. Compare, for example, *I wonder what is/s that up there*, where reduction is possible. In this example there is no deletion site right-adjacent to *is*, and so *is* may be reduced. In (16c), however, the deletion site of *what* is between *is* and *up*, blocking reduction.

9. Another category of deletions different from ellipsed VPs is “pseudogapping,” when an auxiliary is present; such constructions are as bad as the gaps in (27), for example, **Which man did Jay introduce to Ray and which woman did Jim to Tim?*, analogous to (14c), or **Jay wondered what Kay gave to Ray and what did Jim to Tim*, analogous to (14d). Pseudogapping structures are often analyzed very differently (but see Lasnik 1999: chap. 7, which treats them as VP ellipsis plus “remnant raising”), and I avoid them here. A good theory of parsing would show how children make the right selections.

10. Notice that (31a) is well-formed but (9a) is not. In (31a) the null VP is the complement of the adjacent *could've*, but in (9a) it is not the complement of *John's*. This also accounts for the following distinction.

- (i) a. Kim canceled her subscription, and I would've_{VP}e, too.
- b. *Kim canceled her subscription, and I'd've_{VP}e, too.

A null VP following the reduced *'ve* in (ia) is the complement of *would've*, but a null VP following the reduced *'ve* in (ib) is not the complement of *I'd've*, where the reduced auxiliaries have been cliticized to the subject DP and are no longer in the position of Infl with a VP complement.

11. In (ia) the complementizer contained in the clausal complement to *thought* cliticizes to *thought* straightforwardly and may be deleted (unpronounced). In (ib) the lowest *who* cliticizes to *saw*, the intermediate *who* to *think*, and the complementizer to *think + who*, successively like the clitics of (31), again straightforwardly.

- (i) a. I thought [that/0 Ray saw Fay].
 b. Who_i did you think [~~who~~_i that/0 Ray saw ~~who~~_i]?
 c. *Who_i did you think [~~who~~_i that **who**_i saw Fay]?

However, in (ic), the intermediate *who* cliticizes to *think* just like in (ib), but the lowest (boldface) *who* apparently cannot cliticize to *that*, presumably because *that* does not take complements in the usual sense (despite the name “complementizer,” the following clause does not “complete” the meaning of *that* in the way that *Fay* completes the meaning of *saw*) and is not an appropriate host. Likewise for equivalent complementizers in other languages.

Similarly, in (iia) the boldface *who* may not cliticize to the complementizer *how*, because *how* has no complement. Hence the difference with (iib), where each copied element is deleted in the appropriate way, and the sentence is grammatical if not completely felicitous.

- (ii) a. *Who_i do you wonder [~~who~~_i how [**who**_i solved the problem]]?
 b. What_i do you wonder [~~what~~_i how [John solved ~~what~~_i]]?

There is also an interesting range of comparative data resulting from the discomfort in subject DPs with respect to displacement in several languages; see Rizzi 1990: §2.6 and Lightfoot 2006b: §2.5–§2.6 on French, West Flemish, Swedish, and Vata.

Chapter 5

1. This chapter draws on Lightfoot 2017d and my review of Emonds and Faarlund 2014, Lightfoot 2016.
2. The French, of whom Thom is one, see catastrophes frequently, as noted by others; three-syllable French *catastrophes*, /katakastrof/, with stress on the final syllable, seem less catastrophic than four-syllable English /katakastrofiz/ with antepenult syllable stressed. So French and English “catastrophes” have somewhat different meanings as well as different pronunciations.
3. Much of the commentary concerns the genealogy of Middle and Modern English, but if languages are not transmitted in the way presupposed by the cladograms of people defining language “families,” more radical rethinking is needed.
4. Haeberli and Ihsane consider data concerning the distribution of adverbs that has never previously been linked to properties of modal verbs.

