

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)

2 Three New Parses

2.1 That Third Vision

Time now to begin to flesh out the third vision of the previous chapter and to unpack statements there to the effect that all variable I-language properties come from historical change in those I-languages. In this chapter we will reconsider three changes in the I-languages of speakers of English, two of which were addressed in my 1991 effort to show how cue-based children might set parameters (Lightfoot 1991). I addressed the third change in my 1999 book on the development of language, again treating change as a result of parametric shifts. The third vision of chapter 1 takes a very different approach, which we will now begin to work through in detail: we will view children not as setting parameters and evaluating the resulting grammars, counting what grammars do and do not generate, but instead as discovering contrasts and selecting their I-language, postulating structures as allowed by UG and as needed to understand their ambient E-language, their PLD. That is to say, children parse the external language around them, assigning linguistic structures to what they hear from that ambient E-language. Parsing the ambient language is how children discover the variable properties of their emerging I-languages.

I have asserted that any I-language-specific property must have arisen as a result of historical change. I argue that, at least in the domain of syntax, this can happen only through language acquisition by young children. Phonology might be different, but that would be another book, written by somebody else. Here I examine three very specific peculiarities of English, properties that have emerged in English I-languages and not in other European languages. We know enough about when and how they arose to be able to provide good explanations for them in terms of new parses. I will explain what changed and why. Given that very few of the world's languages have a richly recorded history, we will not be able to develop equivalent explanations in a pervasive way, but it will be helpful to think through what it would have taken for I-language peculiarities to have arisen at some point in history. A key to the whole approach is to link changes in E-language with changes in I-languages and to link them by making parsing central to our view of acquisition. As noted earlier, E-language is what is parsed and an I-language results from the parsing, hence there is an intimate relationship between the two.

2.2 Explaining Change through Learnability and Acquisition

In work on syntactic change, one needs good hypotheses for the early stage of the language under investigation and for the late stage, after the change has taken place; one needs a good synchronic analysis both before and after the changes to be considered. That means that one needs all the ideas marshaled by synchronic syntacticians. However, to describe and explain changes through time, questions arise that are not typically explored in synchronic work. Different research strategies are called for, and certainly there are different research traditions involved. Under an approach linking syntactic change to acquisition, work on change casts light on the idealizations used in synchronic work,

as we shall begin to see in §2.4, and is instructive for synchronic syntacticians.

Over recent decades, an approach has developed that links explanation of syntactic changes to ideas about language acquisition, learnability, and the (synchronic) theory of grammar. One way of making this linkage construes change as always externally driven by new PLD.

Baker and McCarthy 1981 identified the “logical problem of language acquisition” as that of identifying the three elements of the following analytical triplet.

(1) Primary linguistic data (Universal Grammar → grammar)

Children are exposed to PLD and as a result their initial state, characterized as UG, develops into a mature state, characterized by a particular, individual grammar, an I-language. The solution to the logical problem lies in identifying the three items in a way that links a particular set of PLD to a particular grammar, given particular ideas about UG. Children seek the simplest and most conservative grammar compatible with both UG and the PLD that they encounter (Snyder 2007).

Under that approach, there can only be one way to explain the emergence of a new grammar. When children are exposed to new PLD that cannot be parsed appropriately by an existing I-language, the new PLD trigger a new grammar. In that sense, it is new PLD that *cause* the change; UG certainly does not cause changes, only providing the outer limits to what kinds of new I-languages may arise. PLD consist of structurally simple things (degree-zero simple, in fact; see §1.1) that children hear frequently, robust elements of their E-language.

Crucial to this approach is Chomsky 1986’s distinction between I- and E-language, both of which play an essential role in explaining change. E-language is the amorphous mass of language out in the world, the things that people hear. There is no system to

E-language; it reflects the output of the I-language systems of many speakers under many different conditions, modulated by the production mechanisms that yield actual expressions. I-languages, on the other hand, are mental *systems* that have grown in the brains of individuals who have parsed their ambient E-language. These systems characterize the linguistic range of those individuals; I-languages are represented in individual brains and are, by hypothesis, biological entities.

New PLD cause change in I-languages. Work in diachrony, therefore, makes crucial use of the E-language–I-language distinction and keys grammatical properties to particular elements of the available PLD in ways that one sees very rarely in work on synchronic syntax. Successful diachronic work distinguishes and links two kinds of changes: changes in PLD (part of E-language) and changes in mature I-languages. These changes are quite different in character, as we will see when we consider our well-understood changes below (§2.4–§2.6) and ask, more broadly, about the role of PLD and whether children are setting parameters or discovering new structures.

Work explaining language change through acquisition by children has been conducted now for decades, and there have been surprising results that lead us to rethink the relationship between PLD and particular I-languages. Diachronic syntacticians have ideas quite different from those common among their synchronic colleagues about which PLD trigger which particular grammars, something that synchronic syntacticians rarely write about.

2.3 Models of Acquisition

Work in synchronic syntax has rarely linked grammatical properties to particular triggering effects, in part because practitioners often resort to a model of language acquisition that is flawed and

strikingly disconnected from work on historical change. I refer to a model that sees children as evaluating grammars globally against sets of sentences and structures, matching input, and evaluating grammars in terms of their overall success in generating the input data most economically (e.g., Clark 1992; Gibson & Wexler 1994, discussed in §1.2). A fundamental problem with this approach is that I-languages generate an infinite number of structures. If children are viewed as setting binary parameters, they must, on the conservative assumption that there are only thirty or forty parameters, entertain and evaluate billions or trillions of grammars, each capable of generating an infinite number of structures (see §1.2).

Beyond these overwhelming issues of feasibility, the evaluation approach raises further problems for thinking about syntactic change, because work often fails to distinguish E-language changes from I-language changes and encounters problems of circularity: the new grammar is most successful in generating the structures of the new system, but in order to explain the emergence of the new grammar, it is presupposed that the new structures are already available. This is part of a larger problem: if one asks a syntactician how children can learn some grammatical property, she will point to sentences that are generated in part through the effects of the relevant grammatical property, taking those sentences to be the necessary PLD. This circularity will become clearer when we discuss specific changes.

New thinking is needed. A “discovery approach” is not subject to the feasibility problems of global grammar evaluation if it treats children as selecting structures expressed by PLD, the structures needed in order to parse sentences. There may be a thousand or more possible structures, but that does not present the feasibility problems of evaluating the success of thirty or forty binary parameter settings against corpora, that is, evaluating the comparative success of complete I-languages with different parameter settings

in generating a given corpus of structures. Children posit structures that are required to analyze what they hear, and that parsing is the key to language acquisition (as suggested by Janet Fodor's important pair of papers: Fodor 1998b,c). Once children have an appropriate set of structures, the resulting I-language generates what it generates and the overall set of structures generated plays no role in triggering or selecting the grammar. That is, children do not perform calculations on what different grammars generate. Children parse, selecting the structures needed to understand what they hear, in principle, one by one in local decisions. A particular grammar is the result, but children do not evaluate the generative capacity of different overall grammars.

This model of acquisition, essentially a discovery procedure in the sense of *Syntactic Structures* (Chomsky 1957; see my introduction to the second edition, Lightfoot 2002), keys elements of grammar to particular elements of the PLD and provides good explanations for diachronic shifts and the emergence of new grammars. Our discovery procedure, however, is complemented by a strong theory of UG, unlike the discovery procedures characterized in *Syntactic Structures*. Under this model, we can link an element of I-language structure with PLD that express that structure, and this yields some surprising results that force us to think about triggering experiences differently.

So a person's internal language capacity is a complex system that depends on an interaction between learned operations and principles that need not be learned but are conveyed by the genetic material, directly and indirectly. The language capacity grows in children in response to the E-language that they encounter, the source of the I-language structures, and becomes part of their mature biology. If language growth in young children is viewed in this way, then we can explain language change over generations of speakers in terms of the dynamics of these complex systems: new I-languages are driven entirely by children responding to new E-language. In par-

ticular, we explain how languages shift in bursts, in a kind of punctuated equilibrium, and we explain the changes without invoking principles of history or ideas about a general directionality or about dispreferred grammars (unlike much modern work on diachronic syntax). The three I-language innovations to be discussed in the next three sections are completely contingent on the new PLD in changing E-language.

Under this approach, there is no separate theory of change, no distinct mechanism dealing just with change. Sometimes there are changes in E-language such that children are exposed to different PLD that trigger a different I-language, as illustrated in §1.3. New I-languages, in turn, yield another new set of PLD for the next generation of children in the speech community. That new E-language, stemming in part from the new I-languages, helps to trigger another new I-language, with further consequences for E-language. If we understand indirect connections in this way, we can explain domino effects in language change (see §2.7 and chapter 5).

This chapter examines a sequence of three reanalyses in the I-languages/grammars of English speakers, three PHASE TRANSITIONS, sets of simultaneous changes that introduced unusual properties not shared by closely related or neighboring languages. In all cases, children are computationally conservative, selecting the simplest I-language consistent with principles of UG and the ambient E-language and PLD (Snyder 2007); in particular, we do not need special principles to weed out “dispreferred” I-languages.

Earlier attempts to formulate a discovery approach to language acquisition (e.g., Dresher 1999; Fodor 1998a; Lightfoot 1999) postulated a set of “cues,” structures provided by UG that children discovered when encountering PLD that needed such structures. That involved postulating rich information at the level of UG, very much against the spirit of the Minimalist Program. In this book, I present the changes quite differently, treating the relevant structures as emerging when children parse the ambient PLD.

Here I show how this works in the new parses emerging in I-languages of English speakers; in the next chapter I will focus more on the novel approach to parsing that I adopt.

2.4 First New Parse: English Modals

Modern English has forms like the (a) examples in (2–6) but not the (b) examples.

- (2) a. He has seen stars.
b. *He has could see stars.
- (3) a. Seeing stars, ...
b. *Canning see stars, ...
- (4) a. He wanted to see stars.
b. *He wanted to can see stars.
- (5) a. He will try to see stars.
b. *He will can see stars.
- (6) a. He understands music.
b. *He can music.

However, earlier forms of English also had the (b) forms; they occur in texts up to the writings of Sir Thomas More in the early sixteenth century. More and some writers before him used all the forms of (2–6), but the (b) forms do not occur in anybody's writing after him. In (7–9) are examples of the latest occurrences of the obsolescent forms, (7) corresponding to (2b), (8) to (4b), and (9) to (5b).

- (7) If wee **had mought** convenient come togyther, ye woulde rather haue chosin to haue harde my minde of mine owne mouthe.

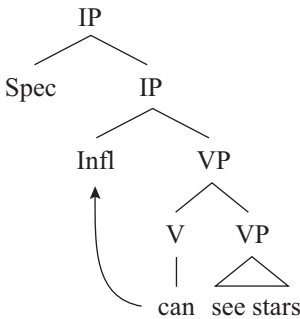
(1528; More, *A Dialogue Concerning Heresies*)

'If we had been able to come together conveniently ...'

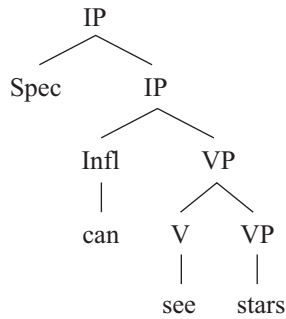
- (8) that appered at the fyrste **to mow** stande the realm in grete stede
(1533; More, *Apology*)
'what appeared at first to be able to stand the realm in good stead'
- (9) I fear that the emperor will depart thence, before my letters **shall may** come unto your grace's hands.
(1532; Cranmer, letter to King Henry VIII)

There is good reason to believe that there was a single change in people's internal systems: *can*, *could*, *must*, *may*, *might*, *will*, *would*, *shall*, *should*, and *do* were previously parsed as more or less normal verbs, but they came to be parsed as Infl (Inflection) or T (Tense) elements. Before More, verbs like *can*, in fact all verbs, MOVED to a higher Infl position, as in (10). After More, verbs like *can* were generated directly as Infl elements and occurred in structures like (11).

(10)



(11)



This single shift in the system was manifested by the simultaneous loss of the (b) forms in (2–6): the phase transition. Sentences like these are not compatible with a system with structures like (11) instead of structures like (10). If perfect and progressive markers

are generated in the specifier of VP, then they will never occur to the left of Infl as in (2b) and (3b). If there is only one Infl in each clause, then (4b) and (5b) will not be generated; (6b) could not be generated by structures like (11), which do not allow Infl to directly precede a DP like *music*.

This change occurred only in Early Modern English and no other language; it was complete by the early sixteenth century. Notwithstanding its uniqueness, positing that this change is somehow indicative of a general tendency does enable researchers to unify it with other phenomena, which offers some degree of explanation. The change of category membership for the English modals is, in fact, a parade case of grammaticalization; but saying that it results from an internal drive, or a general tendency, or a UG bias in that direction gives no explanation for why it happened when it did and under the circumstances under which it did. Nothing similar happened in any other European language, so this change cannot be explained by a “general tendency” to grammaticalize or to recategorize modal verbs as members of a functional category.

A critical property of this change is that it consisted entirely in the *loss* of the (b) phenomena in (2–6), with no new forms emerging. Since children converge on their I-language in response to ambient simple expressions and not in response to negative data about what does not occur, the new, more limited data need to be explained by a new abstract system that fails to generate the (b) phenomena. There were no new forms in which the modal auxiliaries began to occur, so the trigger for the new system must lie elsewhere. In this case, the new PLD cannot be the new output of the new grammars, because there are no new forms. Changes like this, which consist only in the loss of expressions, make a kind of poverty-of-stimulus argument for diachrony: there appear to be no new forms in the PLD that directly trigger the loss of those expressions.

If we ask why this or any other I-language change happened, there can only be one answer under this approach: Children came

to have different PLD as a result of a prior change in E-language. We have a good hypothesis about what the change was in this case.

Early English had complex inflectional morphology. For example, given the inflection of verbs for person and number, we find *fremme, fremst, fremþ, fremmaþ* in the present tense of ‘do’ and *fremed, fremedest, fremede, fremedon* in the past tense; *sēo, siehst, siehþ, sēop* in the present tense of ‘see’; *rīde, rītst, rītt, rīdaþ* for the present tense of ‘ride’ and *rād, ride, rād, ridon* for the past tense. There was a massive loss of verbal morphology in Middle English, beginning in the north of England and due to intimate contact with Scandinavian speakers and widespread English–Norse intermarriage and bilingualism. Again I skip interesting details (see Lightfoot 2017a and §5.4 here), but external language that children heard changed such that the modern modal auxiliaries *can, shall*, and so on came to be morphologically distinct from other verbs. As members of the small preterite-present class, they lacked one surviving feature of person-and-number inflection, the present-tense third-person singular ending *-s*; they had the other surviving feature, the past and present second-person forms in *-st*, but that seems not to have been enough to ensure their survival as forms of verbs. This made them formally distinct from all other verbs, which had the *-s* ending. Furthermore, their “past-tense” forms (*could, would, might*, etc.) had meanings that were not past time, reflecting old subjunctive uses:

(12) They might/could/would leave tomorrow.

The evidence indicates that these modal verbs were parsed differently in people’s internal systems, because they had become formally distinct from other verbs as a result of the radical simplification of morphology (Lightfoot 1999). Thomas More parsed his E-language and had elements like $_{\text{Infl}}[_V \text{can}]$ in his I-language (a verb *can* moved to an Infl position), while after him speakers had $_{\text{Infl}}\text{can}$, where *can* was no longer parsed as a verb. So we see domino

effects: changes in what children heard, the newly reduced verb morphology, led to a different categorization of certain verbs, which yielded systems like (11) that were compatible with the (a) forms of (2–6) but not with the (b) forms.

Thomas More is the last known person with the old system. For a period, both systems coexisted: some speakers had (10) and others had (11), the former becoming rarer over time, the latter more numerous. A large literature is now devoted to this kind of sociological variation, changing over time, and we will return to this matter in §2.7.

Parsing, of course, depends on contrasts: in particular, formal and distributional contrasts influence the categorization of words. Words like *kick*, *like*, and *seem* are formally marked for person, number, and tense, co-occur with adverbs like *often* and *tomorrow*, and are categorized as verbs. *Girl*, *caterpillar*, and *catastrophe* are formally marked for number but not person or tense, co-occur with determiners like *the* and adjectives like *tall* and *unbelievable*, and are categorized as nouns. As a result of the simplification of morphology, verbs became definable as having the *-s* ending. In that case, *can* was no longer parsable as a verb, was categorized differently, and as a result developed a new distribution. This all makes sense, assuming uncontroversially that words are parsed as members of syntactic categories and that the parses might change over time. However, it is hard, not to say impossible, to see how the new variable properties could be seen as manifesting a new setting of a binary and UG-defined structural parameter, despite Lightfoot 1991.

2.5 Second New Parse: Verbs Ceasing to Move

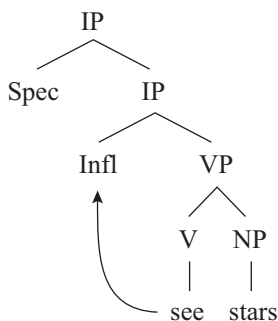
A later major change was that English lost forms like the (a) forms in (13–15), another phase transition of simultaneous disappearances. Such forms occur frequently in texts up through the seventeenth

century, diminishing over a long period in favor of the (b) forms and ultimately disappearing.

- (13) a. *Sees Kim stars?
 b. Does Kim see stars?
- (14) a. *Kim sees not stars.
 b. Kim does not see stars.
- (15) a. *Kim sees always stars.
 b. Kim always sees stars.

Again we can understand the parallelism of the changes in terms of a single change in the abstract system, namely the loss of the operation moving verbs to a higher Infl position:

(16)



This is another change that did not affect other European languages, whose systems have retained the verb-movement operation (apart from Faroese and, perhaps, some Scandinavian systems; see Heycock et al. 2012 and references therein). Present-day English verbs do not move to the higher position and therefore cannot move to a clause-initial position (13a), to the left of a negative (14a), or to the left of an adverb (15a). The equivalent movements continue to occur

in French, Italian, Spanish, Dutch, and German systems. Again a contingent explanation is required: what was it about English at this time that led to this shift in I-languages? In particular, what were the new PLD that children were exposed to that helped to trigger the new parse?

It is plausible that this shift was due to two prior changes in E-language and that we see here another domino effect, a sequence of changes. The first change was the output of the new I-language that we just discussed, involving the new parsing of a distinct category of modal verbs, which are very frequent in typical speech (see Leech 2003). Given that words like *can* and *must* were no longer verbs but Inflection items, no sentence containing such a word would have a $_{\text{Infl}}\text{V}$ structure.

The second change was the emergence of “periphrastic” *do* forms as an alternative option for expressing past tense: *John did leave*, *John did not leave* instead of *John left* and *John left not*. Given that *do* forms were instances of Inflection, any sentence containing one would not have the $_{\text{Infl}}\text{V}$ structure. As a result of these changes, beginning on a large scale in the fourteenth century, the Infl position came to be heavily occupied by modal auxiliaries and *do*. Thus, lexical verbs did not occur in the Infl position as often as before the days of periphrastic *do* and before modal auxiliaries were no longer verbs, and as a result the $_{\text{Infl}}\text{V}$ structure was expressed much less. Apparently it fell below the threshold that had permitted its selection by children. The loss of the (a) forms in (13–15)—a loss that reached totality only in the eighteenth century—suggests that a new system emerged in which the Infl position was no longer available as a target for verb movement. The conservative Thomas More had parsed his E-language to have $_{\text{Infl}}\text{V}$ structures, reflecting the movement of verbs to Infl; after the eighteenth century individuals had no such structures but instead had $\sqrt{[\text{V} + \text{Infl}]}$, reflecting the attachment of inflectional endings onto a lower, unmoved verb

(subject to Howard Lasnik's Stranded-Affix Filter, ensuring that affixes are attached somewhere: Lasnik 2000: 123).

As with the first new parse, the two systems coexisted for a while, in fact for a longer period in this case: Shakespeare and other writers alternated easily between the coexisting old and new systems, sometimes using the old V-to-Infl forms and sometimes the new *do* forms, even in adjacent sentences, as in the following examples from Shakespeare's *Othello*.

- (17) a. Where **didst thou** see her?—O unhappy girl!—With the Moor, **say'st thou**?
 b. I **like not** that.—What **dost thou** say?
 c. Alas, what **does this gentleman** conceive? How **do you**, madam?

Again this is too brief an account (see Lightfoot 2017a), but it is clear that prior changes in E-language, some due to a shift in I-languages, had the effect of reducing enormously children's evidence for the $_{\text{Infl}}\text{V}$ structure, triggering a new internal system and a new parse. These simultaneous but apparently unrelated changes were a function of that single change in the abstract system, the loss of $_{\text{Infl}}\text{V}$ structures, a genuine phase transition consisting of diverse but simultaneous changes.¹

2.6 Third New Parse: Atoms of *Be*

There is a third, remarkable phase transition, observed and analyzed in Warner 1995, which results in part from the two changes just discussed. It involves very peculiar properties of the verb *be*, which have no equivalent in other European languages that are closely related to English. It certainly lies far beyond any viable theory of parameters. One way of characterizing the change is that different forms of the verb *be* came to be listed in the mental lexicon as

atomic, or “monomorphemic” as Warner puts it, and developed their own subcategorization frames.

First, some background relating to VP ellipsis. VP ellipsis is generally insensitive to morphology and one finds cases where the understood form of the missing verb differs from the form of the antecedent:

- (18) a. Kim **slept** well, and Jim will [sc. **sleep** well] too.
 b. Kim **seems** well-behaved today, and she often has [sc. **seemed** well-behaved] in the past, too.
 c. Although Kim **went** to the store, Jim didn't [sc. **go** to the store].

There is a kind of SLOPPY IDENTITY at work here, since *slept* and *sleep* in (18a) are not strictly identical but “sloppily” so. Similarly in (18b) and (18c). One way of thinking of this is that in (18a) *slept* is analyzed as [_v*sleep* + past] and the understood verb of the second conjunct accesses the verb *sleep*, ignoring the tense element.

However, Warner noticed that *be* works differently: it occurs in elliptical constructions only on condition of STRICT IDENTITY with the antecedent. In (19a,b) the understood form is strictly identical to the antecedent, unlike in the nonoccurring (19c–e).

- (19) a. Kim will **be** here, and Jim will [sc. **be** here] too.
 b. Kim has **been** here, and Jim has [sc. **been** here] too.
 c. *Kim **was** here and Jim will [sc. **be** here] too.
 d. *If Kim **is** well-behaved today, then Jim probably will [sc. **be** well-behaved] tomorrow.
 e. *Kim **was** here yesterday and Jim has [sc. **been** here] today.

This suggests that *was* is not analyzed as [_v*be* + past], analogously to *slept*, and that forms of *be* may be used as an understood form only when precisely the same form is available as an antecedent, as in (19a,b).

Warner notes that the ellipsis facts of modern English *be* were not always so, and one finds forms like (19c–e) in earlier times. Jane Austen was one of the last writers to use such forms: she used them in her letters and in dialogue in her novels, as for example in (20a,b), but not in narrative prose, presumably indicating a conversational style. These forms also occur in eighteenth-century writings, such as in (20c), and earlier, when verbs still moved to Infl, as in (20d).

- (20) a. I wish our opinions were the same. But in time they will
[sc. be the same].
(1816; Jane Austen, *Emma*, ed. R. W. Chapman [London: Oxford University Press, 1933], p. 471)
- b. And Lady Middleton, is she angry? I cannot suppose it possible that she should [sc. be angry].
(1811; Jane Austen, *Sense and Sensibility*, ed. R. W. Chapman [London: Oxford University Press, 1923], p. 272)
- c. I think, added he, all the Charges attending it, and the Trouble you had, were defray'd by my Attorney: I ordered that they should [sc. be defrayed].
(1741; Samuel Richardson, *Pamela* [London, 3rd ed.], vol. 2, p. 129)
- d. That bettre loved is noon, ne never schal.
(c1370; Chaucer, *A complaint to his lady*, line 80)
'So that no one is better loved, or ever shall [sc. be].'

These forms may be explained by supposing that, in (20a) for example, *were* is analyzed as [_v*be* + subjunctive] and the *be* is accessed by the understood *be* in the following *but* clause. That is, up until the early nineteenth century the finite forms of *be* are decomposable, just like ordinary verbs like *sleep* in present-day English. Hence the loss of expressions like those of (20) would be attributed to the new, monomorphemic parsing of the *be* forms.

Warner goes on to show that, now that forms of *be* are atomic and undecomposable, present-day English shows quite idiosyncratic

restrictions on particular forms of *be*, which did not exist before the late eighteenth century or early nineteenth century. For example, it is only the finite forms of *be* that may be followed by a *to* infinitive, as shown by (21); only *been* may occur with a directional preposition phrase, as (22) shows; and *being* is subcategorized as not permitting an *-ing* complement, as in (23).

- (21) a. Kim was to go to Paris.
 b. *Kim will be to go to Paris.
- (22) a. Kim has been to Paris.
 b. *Kim was to Paris.
- (23) a. I regretted Kim reading that chapter.
 b. I regretted that Kim was reading that chapter.
 c. *I regretted Kim being reading that chapter.

Restrictions of this type are stated in the lexicon. These idiosyncrasies show clearly that *been*, *being*, and so on must be listed as individual lexical entries in order to carry their own individual subcategorization restriction. However, these restrictions did not exist earlier and one finds forms corresponding to the nonoccurring sentences of (21–23) through the eighteenth century: (24a) is equivalent to (21b), (24b) to (22b), and (24c) to (23c).

- (24) a. You will be to visit me in prison with a basket of provisions.
 (1814; Jane Austen, *Mansfield Park*, ed. J. Lucas [London: Oxford University Press, 1970], p. 122)
- b. I was this morning to buy silk.
 (1762; Oliver Goldsmith, *Citizen of the World*)
 Meaning: ‘I went to ...’, not ‘I had to ...’
- c. Two large wax candles were also set on another table, the ladies being going to cards.
 (1726; Daniel Defoe, *The Political History of the Devil* [Oxford: Talboys, 1840], p. 336)

So there were changes in the late eighteenth century to early nineteenth century whereby the ellipsis possibilities for forms of *be* became more restricted and particular forms of *be* developed their own idiosyncratic subcategorization restrictions, both properties indicating the new, undecomposable, monomorphemic nature of the forms of *be*.

I-languages allow computational operations on items stored in a mental lexicon, and both the operations and the items stored may change over time. There is good reason to believe that decomposable items like [_V*be* + subjunctive] and [_V*be* + past] ceased to be stored in that form, replaced by undecomposed, atomic forms like *were*, *was*, *been*, each with its own subcategorization restrictions. The phenomenon has no parallel in closely related languages, but we can explain it by showing how it may have arisen in the I-languages of English speakers at this time.

It is natural to view this change as a consequence of the changes discussed in §2.4 and §2.5. After the loss of rich verb morphology and the loss of the _{Infl}V structures, the category membership of forms of *be* became opaque, leading to new structures being assigned through new parses. If the *be* forms were instances of V, then why could they occur where verbs generally cannot occur, for example, to the left of a negative or, even higher, to the left of the subject DP (*She is not here; Is she happy?*)? If they were instances of Infl, then why could they occur with another Infl element such as *to* or *will* (*I want to be happy; She will be here*)?

In earlier English, forms of *be* had the same distribution as normal verbs. After the two phase transitions discussed earlier, they had neither the distribution of verbs nor that of Infl items. The evidence is that from the late eighteenth century, children developed I-languages that reflected new parsing, treating forms of *be* as verbs that have the unique property of moving to higher functional positions and being undecomposed, atomic elements, unlike other verbs. It is impossible to see how such specific variable properties might

be captured by a binary parameter defined at UG. However, a child parsing relevant structures might indeed select the new parses that are required.

2.7 Domino Effects

Historical changes in English I-languages can be understood as the result of children acquiring their I-language as their PLD change. We have seen examples of phase transitions, when several phenomena change simultaneously. We also see domino effects, when a number of phenomena change not simultaneously but in rapid sequence. For example, English underwent massive simplification of its verb morphology, initially under conditions of bilingualism in the northeast of England (see §5.4 for more on the Scandinavian character of Middle English). The new PLD led to a new I-language with a dozen former verbs now parsed as Infl items, categorized as instances of Infl. As a result, the PLD changed again; combined with new periphrastic forms with *do*, this led to new I-languages where verbs ceased moving to higher Infl positions. This, in turn, led to new PLD in which the categorial status of forms of *be* became opaque, leading to the reanalysis of §2.6 and revealing a sequence of new parses.

The modern distinction between E-language and I-language is crucial to this analysis; both contribute to explaining change (Lightfoot 2006a). We see a dynamic interplay between changes in E-language (new PLD) and changes in I-languages. New E-language leads to new I-languages, new I-languages lead to new E-language, and sometimes we see sequences of changes, domino effects, which we can understand if a central component of language acquisition is the parsing of E-language. In the three case studies examined in the last three sections, we see causal relationships between E-language and I-language changes and the particular contingencies that triggered new I-languages at particular times.

When a new I-language, I-language_p, develops in one individual, that changes the ambient E-language for others, making it more likely that another child will acquire I-language_p; likewise for the next child, and so on. As a result, the new I-language spreads through the speech community quickly. Niyogi 2006 provides a computational model of how new language systems might spread quickly through speech communities; see §5.9.

2.8 Variable Properties

Children are exposed to speech, and their biological endowment, a kind of toolbox, enables them to begin to parse their external linguistic experience, thereby discovering I-language elements and growing a private, internal system that defines their linguistic capacity. Internal systems involve particular abstractions, categories, and operations, and these constitute the real points of variation and change; children need to discover them through parsing. Phenomena do not always change in isolation but often cluster, depending on the abstract categories involved. As a result, change is bumpy and takes place in “punctuated” bursts that disrupt general equilibrium. We explain the bumps, the clusters of changes, in terms of changes in the abstract system; the three phase transitions in the history of English that we have looked at provide an illustration. If we get the abstractions right, we explain why phenomena cluster as they do. More on this in chapter 5; meanwhile the three analyses of this chapter seem to explain the observed changes via changes in the abstract system, without needing to invoke parameters.

Everybody’s experience varies and people’s internal systems may vary, but not in any simple, linear fashion. I-languages change over time, and sometimes variation in E-language experience (new PLD) is sufficient to trigger the growth of a different internal system. Children are sensitive to variation in E-language, to variation in

initial conditions, in the terminology of chaos theory, and this influences how they parse expressions. For example, after the comprehensive morphological changes of Middle English, young children had different experiences that led them to categorize words like *may* and *must* differently from verbs like *run* and *talk*. The assignment of these words to a different category, Infl (or T), explains why the (b) structures in (2–6) all disappeared in parallel. Similarly, new structures resulting from modal verbs being treated as Inflection items and new structures with periphrastic *do* entailed that the $_{\text{Infl}}V$ structure was expressed much less robustly and fell out of use, entailing the obsolescence of the (a) versions of (13–15). Finally, as a result of these two phase transitions, forms of *be* were parsed as monomorphemic and no longer as a verb plus a separate inflectional affix.

Under this approach, change and therefore variability is contingent, dependent on particular circumstances, which accounts for why English at this time underwent changes that other European languages have not undergone at any point. English had particular morphological properties that were affected in particular ways by contact with Norse speakers and that led to the new categorization, the new parses. If change is contingent like this, then there is no general direction to change and no reason to believe that languages all tend to become more efficient, less complex, and so on. There are no general principles of history of the kind that nineteenth-century thinkers sought (e.g., Darwinians seeing “progress” in new species and Marxists seeing certain kinds of societies changing into other specific types) and that modern diachronic syntacticians continue to invoke. Explanations are local (Lightfoot 2013) and there is no reason to revive historicism, or declare principles of history, or UG biases. There is no reason to invoke UG biases or to expect that variable properties in I-languages will fall into a narrow class defined by a restrictive theory of

parameters, as people following the second vision of chapter 1 imagine.

This approach to syntactic change also provides a new understanding of synchronic variation, along the lines of William Labov's 1972 discussion of sound change in progress. When a phase transition takes place, it does not happen on one day, with all speakers changing in unison. Rather, a new I-language emerges in some children and spreads through the population, sometimes over the course of a century or more but usually not for a very long period (although see the very interesting Wallenberg 2016 on the slow loss of extraposition structures). Competing grammars (Kroch 1989) explain the nature of certain variation within a speech community: in this context, one does not find random variation in the texts but oscillation between two (or more) stable organizations, two I-languages. In general, writers either have all the forms of the obsolescent I-language or none. Not all variation between texts, of course, either by the same or different writers, is to be explained in this way: only variation in I-languages.

There is also variation in E-language that has little if anything to do with I-language. E-language is amorphous, and in it variation is endemic and does not come in the structured form that variation in I-languages shows. No two people experience the same E-language, and in particular, no two children experience the same PLD. Since E-language varies so much, there are always possibilities for new I-languages to be triggered. New I-languages may come to incorporate unusual and very particular properties, as we have seen and as we shall see again when we discuss the nature of parsing in §3.1 and the spread of I-languages in §5.8. We see a wider range of variable properties than the parametric approach leads us to expect, and we understand how new variable properties emerge by seeing them as emerging through the parsing of new E-language.

2.9 Identifying Triggers

The major contributions of diachronic work in syntax lie in explaining one kind of variation, the kind due to coexisting I-languages, and in revealing, for any particular property of I-languages, what the E-language trigger might be.

It is surprising how little discussion there has been among synchronic syntacticians of what triggers what properties, given the generally accepted explanatory schema of (1), the “analytical triplet” of Baker and McCarthy 1981. Reducing hypothesis space is an essential part of the enterprise but not sufficient: we also want to know what aspects of new E-language might have triggered new I-languages.²

We should allow for the possibility that the PLD that trigger a particular parse may not have any obvious connection with that structure. Indeed, the recategorization of modal verbs was triggered by new morphological properties.

Niko Tinbergen (1957: chap. 22) once surprised the world of ethologists by showing that young herring gulls’ behavior of pecking at their mothers’ beaks was triggered not by the fact that the mother gull might be carrying food but by a red spot that mother gulls typically have under their beak. Tinbergen devised ingenious experiments showing that the red spot was the crucial triggering property. Chicks would respond to a disembodied red spot but not to a bird carrying food with the red spot hidden. Similarly and as we saw in our first case study, grammars may have a given mechanism or device as a result of properties in PLD that are not obviously related to that mechanism.

Furthermore, there is no reason to believe that elements of different I-languages are always triggered by the same PLD. For example, children developing an English I-language could learn that VPs have verb–complement order from a simple sentence like *Bent visited Oslo*. Since verbs do not move in English I-languages and

Infl lowers on to verbs, *visited Oslo* can only be analyzed as $_{VP}[visit + Infl\ Oslo]$, verb–complement. A Norwegian child, however, could not draw the same conclusion from the word-for-word translation *Bent besøkte Oslo*, which does not reveal the underlying position of the verb. Norwegian I-languages are verb-second: finite verbs move to a high, “second” position (presumably in the C projection), yielding simple structures like $_{CP}[Bent\ C\ besøkte\ IP[Bent\ I\ besøkte\ VP[besøkte\ Oslo]]]$. The verb-second analysis is required by synonymous sentences like *Oslo besøkte Bent*, where the finite verb surfaces to the left of the subject DP and requires a structure $_{CP}[Oslo\ C\ besøkte\ IP[Bent\ VP[besøkte\ Oslo]]]$, as well as “topicalized” expressions like $_{CP}[Søndager\ C\ besøkte\ IP[Bent\ I\ besøkte\ VP[besøkte\ Oslo]]]$ ‘On Sundays Bent visited Oslo’. Therefore, *Bent besøkte Oslo* does not reveal the structure of the VP, and a more complex sentence like *Bent kan besøke Oslo* ‘Bent can visit Oslo’ is needed to express $_{VP}[V\ DP]$ structures, *Bent kan* $_{VP}[besøke\ Oslo]$. Similarly, in German, another verb-second language but with complement–verb order, the complex *Bent kann Oslo besuchen* reveals the $_{VP}[DP\ V]$ structure.³

Speakers have their own internal system, an I-language, that grows in them in the first few years of life as a result of an interaction between genetic factors common to the species and environmental variation in PLD. Such a grammar represents the person’s linguistic range, the kind of things they might say and how they may say them. If children hear different things, they may converge on a different system, perhaps the first instance of a new I-language. We want to find out what triggers which aspect of a person’s I-language, to understand how new I-languages might emerge.

If we can discover things about I-languages by looking at how they change, we can generate productive hypotheses about what PLD trigger particular properties of I-languages, thereby explaining the I-language properties. That will surely interest synchronic syntacticians.

Theodosius Dobzhansky (1964: 449) noted famously that “nothing in biology makes sense except in the context of evolution,” later the title of his famous 1973 paper in *American Biology Teacher*. His position was that biologists can explain properties of organisms by showing how they might have arisen through evolution. This stance is also taken by Minimalists who argue that the rich information postulated in Government and Binding approaches to UG is evolutionarily implausible. Here I have shown that there is a parallel line of argument, explaining how new language-specific properties—but not properties of UG, of course—evolve through historical change. We may have achieved ideal explanations for certain syntactic changes in terms of how children select their I-language.

2.10 Conclusion

By identifying shifts in the ambient E-language that plausibly triggered new I-languages, we can explain diachronic changes in I-languages in terms of language acquisition, distinguishing I-languages and E-language, but with each playing a crucial role. This provides a model for explaining other unusual properties that occur in mature syntactic systems. Parsing is key, implemented by the individual’s I-language. Parameters and evaluation metrics play no role.

Linking matters of acquisition and learnability to matters of syntactic change permits deep explanations of particular changes and illuminates what experience it takes to trigger particular elements of I-languages. Under this approach, there is no independent theory of change, and instead change is an epiphenomenon. Children acquire their own private I-language when exposed to the ambient E-language, not influenced directly by any ambient I-language, which cannot be observed. No two children experience identical E-language. Therefore, there is always the possibility of different I-languages emerging, but nothing is actually transmitted and there

is no object that changes. Rather, different I-languages emerge in different children, and I-languages with certain properties may spread through a language community, through the medium of E-language.

We explain peculiarities in I-languages by showing how they arise through language acquisition taking place when E-language changes for a generation of children. We therefore explain the peculiarities in ways that are broadly similar to the ways in which biologists explain properties of organisms by showing how they might have evolved in the species (see chapter 6). All I-language variable properties fall within a narrow range, permitted by the Merge and Project structure building allowed by the invariant principles of UG. Where we have insufficient data to tell rich stories like the ones documented here, we should nonetheless be able to imagine plausible scenarios, just as the imagination of biologists was provoked by Dobzhansky.

