STiL
Studies in Linguistics

Special Issue on
“Directionality of Phrase Structure Building”

Edited by:
Cristiano Chesi

Studies in linguistics
(Centro interdipartimentale di studi cognitivi sul linguaggio)
ISSN 2281-3128

Correspondence can be addressed to:
CISCL – Centro Interdipartimentale di Studi Cognitivi sul Linguaggio
Dipartimento di Scienze della Comunicazione
Complesso S. Niccolò, Via Roma, 56
I-53100 Siena, Italy

or by email at:
c.chesi@unisi.it
Contents

Cristiano Chesi
Introduction to directionality 7

Colin Phillips and Shevaun Lewis
Derivational Order in Syntax: Evidence and Architectural Consequences 11

Ruth Kempson, Ronnie Cann, And Lutz Marten
Treegrowth Dynamics 49

Ash Asudeh
Directionality and the Production of Ungrammatical Sentences 83

Cristiano Chesi
Do the ‘right’ move 107

Jieun Kiaer
Why Left-to-Right Grammar? Evidence from Korean 139

Michael Brody
Syntax and Symmetry 167
Introduction to directionality

CRISTIANO CHESI
NETS – IUSS Pavia
cristiano.chesi@iusspavia.it

A grammar that uses (recursive) structure building operations is adequate if and only if it recognizes and generates nothing but interpretable, well-formed sentences that are part of the language that the grammar aims at characterizing. If structure building operations overgenerate by recognizing and generating sentences that are considered ungrammatical by native speakers of the language under analysis, some relevant constraint should be added to the grammatical description. This necessity was clearly stated in Ross (1967), and its urgency was revealed by operations like Move α before (Lasnik and Saito 1992), then Merge, Move and Agree more recently (Chomsky 1995-2012): these operations are very general and simply too powerful. On the one hand, they might highlight important universal principles of human language structure, but, on the other, they require too many restrictions to meet, ultimately, the upper bound of the relevant linguistic description. It is a matter of fact that these operations are often defined in such a way to be practically insufficient to circumscribe many relevant empirical facts. Furthermore, common constraints (e.g. binary merge, locality or derivation by phase...) that have been proposed to limit the generative power of grammars are not explicitly related to notions of complexity or simplicity in any explicit computational sense. For instance, neither time or space complexity measures are fully taken into consideration nor a formalization of the problem space is provided (cf. Barton et al. 1987); in the end, what might seem a simpler structure building operation lead to rather complex restrictions in order to attain to empirical (descriptive) adequacy.

In this special issue, nine authors investigated the empirical fit of “simpler” grammatical theories, discussing how certain restrictions on structure building procedures, or different formulations of them, lead to theories that are more descriptively and explanatory adequate. This is done in this volume by mainly focusing on how linear constraints restrict, or are restricted by, the hierarchical structure that is built incrementally.

In this vein, the arguments presented in this volume discuss explicit formalizations of grammars that are largely derivational and directional: derivational since constituents and dependencies are built dynamically, piecemeal, using structure building operations like Merge, Move (Minimalism, Chomsky 1995), or update (Dynamic Syntax, DS, Cann et al. 2005), or monotonic integration of subsequent fragments, (Lexical Functional Grammar, LFG, Bresnan 2001); directional, since structure building operates strictly from left to right as proposed by Colin Phillips (1996, 2003, Phillips and Lewis in this volume, Kempson, Cann, Marteen, and Kiaer in this volume) and top-down (Chesi 2004, 2012, Chesi, Asudeh, and Brody in this volume) Vs. bottom-to-top (Chomsky 1995).
The evidence discussed here is both formal and empirical and it is mostly related to a revised definition of derivation, explicitly addressing foundational issues, like the opportunity of such directionality specification as related to some constrained relation between hierarchical and linear (temporal) order (cf. Kayne 1994), or the necessity of a phase-based (chunked) incremental derivation. These considerations brought attention to empirical data that go beyond the classic performance and competence divide: in which (formal) sense parsing and competence are related? Is it really useful to assume that the grammar specification is really “implementation independent”?

Colin Phillips and Shevaun Lewis (“Derivational Order in Syntax: Evidence and Architectural Consequences”), for instance, argue in favor of a grammatical theory that should be “implementation dependent”: the grammar should operate incrementally, building phrase structures consistently as real-time processing tasks (comprehension and production) demand, namely, roughly “from left to right”. In their lucid review, three main positions, relating the standard bottom-to-top (Minimalist) approach with the mental reality of the grammar are discussed: the literalist one (“bottom-to-top 'cyclic' derivations may be understood as literal descriptions of sequences of mental operations that speakers may, in principle, carry out”), the formalist (“the sequence of steps in a grammatical derivation may be understood not as a temporal ordering of structure-building operations, but rather as a formal relation between a set of structural representations”) and the extentionalist one (“a grammar is merely an abstract characterization of a function whose extension is all and only the well-formed sentences of a given language”) eventually arguing against the elusive nature of the third, widely spread, position. In the end, this paper argues in favor of a closer correspondence between grammatical derivations and the mental processes involved in producing or understanding sentences in real time: this conclusion supports a Left-Right derivational model (Phillips 1996) and, despite apparent problematic empirical issues, is both psycholinguistically tenable and empirically superior.

This is also the conclusion reached by Ruth Kempson, Ronnie Cann and Lutz Marten (“Treegrowth dynamics”). In their paper, these authors provide evidence in favor of a novel grammar formalism (DS), that models natural languages as mechanisms for incrementally building up context-relative semantic representations following the dynamics of the parse process. The point of departure for this framework is to take the concepts of underspecification and update which are familiar in pragmatics and semantics, and extend them to syntax. Some empirical evidence comes from ordering effects (e.g. scrambling). Their work, here, focus on how the morphological puzzle of the rigid relative ordering of clitic pronouns in the Romance languages (Person Case Constraint, PCC) could be explained as a lexical calcification of scrambling mechanisms, as a frozen set of reflexes of what had been freely available in earlier stages of the language under analysis.

Within a different formal framework (Lexical Functional Grammar), Ash Asudeh (“Directionality and the production of ungrammatical sentences”) embraces a similar directionality perspective. Focusing on production contexts in which sentences that are considered “ungrammatical” are commonly produced, a clear asymmetry emerges suggesting the superiority of a Top-Down derivation. This is the case of resumptive pronouns in certain structures: some English speakers produce resumptive pronouns
under certain circumstances, yet reject sentences containing resumptives as ungrammatical. Here Asudeh argues in favor of a language production model that explains this puzzle by assuming that the sentence derivation is incremental and prioritizes local well-formedness over global well-formedness. This has implication also for the directionality of the derivational process: in English resumptive in islands imply that the unbounded dependencies in grammar are filler-driven, featurally defined at the top of the dependency, and that islands are identifiable by the grammar from outside the island, in a top-down fashion (this perspective is also shared with Chesi in this volume). The intuition is formalized within the LFG framework by implementing a production model which is incremental and based on monotonic information growth. LFG provides a formal, ‘outside-in’ theory of unbounded dependencies that treats them as filler-driven and allows island barriers to be identified from outside of the island, rather than from the inside. Outside-in uncertainty presents a theoretically and empirically attractive alternative which has no clear analogue in other standard minimalist theories. Asudeh effectively pursues the outside-in (top-down, left-to-right) approach, presenting some of its theoretical and empirical virtues, and assessing certain challenges that have previously been thought to argue for an inside-out account (e.g. minimalism).

Cristiano Chesi (“do the ‘right’ move”) as well, reaches the same conclusion looking at the basic “filters” commonly operating on free-merge: a major restriction, such as categorial selection, logically leads to a derivation that proceeds Top-Down; if selection is interpreted as an expectation that must be satisfied locally, what is expected must follow what creates the expectation, hence a “left-right” derivation is predicted. The right-branching tendency of such derivation competes with the possibility of expanding an expectation introducing extra (unexpected but compatible) features piecemeal, by specifying or lexicalizing the expected ones. This intuition is implemented within a Phase-based (Chesi 2012) Minimalist framework (Stabler 1997), i.e. a derivational approach where a phase is the minimal expansion domain within which a given expectation must be fulfilled. As a consequence, a partition can be made, distinguishing “true” recursion (phases that are expanded while the originating phase is not yet complete, i.e. nesting) from “tail” recursion (phases computationally independent from the originating one, i.e. sequential) (Abelson & Sussman 1996). Such division explain classic A and A’ movement constraints (e.g. locality and islandhood) by implementing filler-gap, filler-first (Fodor 1978) dependencies as memory-buffered, “rightward” dependencies. Even though “rightward movements”, like Extraposition and Heavy NP-shift, which are discussed thoroughly here, have the same directionality, the relevant empirical constraints, distinguishing classic “leftward” from “rightward” dependencies can be easily captured, e.g. recasting the right-roof constraint in terms of phasal nesting.

Jieun Joe Kiaer (“Why Left-to-Right Grammar?: Evidence from Korean”) presents other evidence for incremental, left-right structure building in Korean, by showing asymmetries and preference for early association among other facts. This evidence supports a derivation (both in comprehension and production) that is driven by structural optimization needs, which are expressed as a strong tendency to build any meaningful, communicative proposition as quickly as possible with a minimized structure-building effort (cf. Hawkins, 2004). Though the nature of discussion is
theory-neutral, to formalize the procedural competence here described, the DS framework (cf. Kempson, Cann & Marten in this volume) is used. Word-order variation in Korean (i.e. scrambling) brings new insight into the study of verb-final languages and also into the architecture of grammar, including the competence-performance relation. To support the claim of step-by-step, left-right structure building, results from perception and production tests are discussed, which show that Korean native speakers strongly prefer to resolve any grammatical dependency within the ‘first’ available slot before the whole structure becomes visible.

**Michael Brody**’s contribution (“Syntax and Symmetry”) concludes this collection by approaching the (explanatory) adequacy of the grammatical theory by investigating possible logical consequences of the Mirror Theory (Brody 2000). Within this representational framework, minimalist, antisymmetrical, structure building operations like labelled Merge, Spec-Head, Head-Complement, Probe-Goal and C-command dependencies are discussed, eventually concluding that one single kind of linear ordering is possible, that is a consequence of the domination relation representing the “top–down” antisymmetrical structure of the tree. This intuition is exploited by interpretive labelling (cf. "super-saliency" idea discussed by Schlenker, 2004, to explain Principle-C effects) and this seems to be the common basis for both (morpho-) phonological linearization and (quasi-) semantic c-command phenomena.

In the end, I believe that the empirical and formal evidence discussed in this volume, bridging the gap between purely formal grammars and psycholinguistic findings can hardly be ignored.

**References**


Derivational Order in Syntax: Evidence and Architectural Consequences*

COLIN PHILLIPS & SHEVAUN LEWIS
University of Maryland
{colin, shevaun}@umd.edu

Most formal syntactic theories propose either that syntactic representations are the product of a derivation that assembles words and phrases in a bottom-to-top and typically right-to-left fashion, or that they are not constructed in an ordered fashion. Both of these views contrast with the (roughly) left-to-right order of structure assembly in language use, and with some recent claims that syntactic derivations and real-time structure-building are essentially the same. In this article we discuss the mentalistic commitments of standard syntactic theories, distinguishing literalist, formalist, and extensionalist views of syntactic derivations. We argue that existing evidence favors the view that human grammatical representations are the product of an implementation dependent system, i.e., syntactic representations are assembled in a consistent order, as claimed by grammatical models that are closely aligned with real-time processes. We discuss the evidence for left-to-right syntactic derivations, and respond to critiques of a proposal that the conflicts between the results of constituency diagnostics can be explained in terms of timing.

1. Introduction
Standard generative grammars describe language in terms that appear distant from considerations of everyday, real-time language processes. To some this is a critical flaw, while to others this is a clear virtue. One type of generative grammar defines a well-formed sentence as a static, structured representation that simultaneously satisfies all relevant constraints of the language, with no regard to how the representation is assembled (e.g., Sag, Wasow, & Bender, 2003). Another type of generative grammar defines a well-formed sentence as a derivation, or sequence of representations, that describes how the sentence is gradually assembled, often including various transformations that move words or phrases from one position to another in a structure. In the most popular current version of the derivational approach, derivations proceed ‘upwards’, starting from the most deeply embedded terminal elements in the sentence, which are often towards the right of a sentence (e.g., Chomsky, 1995; Carnie, 2006). Such derivations tend to proceed in a right-to-left order, which is

* We would like to thank Jeff Lidz, Norbert Hornstein, Terje Lohndal, Ewan Dunbar, and Bradley Larson for useful discussions of the material in this paper. Preparation of the paper was supported in part by NSF grants BCS-0848554 and DGE-0801465 to CP and by a University of Maryland Flagship Fellowship to SL.
probably the opposite of the order in which sentences are assembled in everyday tasks such as speaking and understanding. Since these theories make no claim to being accounts of such everyday processes, the discrepancy causes little concern among the theories’ creators. Generative grammars are typically framed as theories of speakers’ task-independent knowledge of their language, and these are understood to be distinct from theories of how specific communicative tasks might put that knowledge to use.

Set against this background are a number of more recent proposals that various linguistic phenomena can be better understood in terms of derivations that incrementally assemble structures in a (roughly) left-to-right order. One can evaluate these proposals based simply on how well they capture the acceptability judgments that they aim to explain, i.e., standard conditions of ‘descriptive adequacy’. But it is hard to avoid the question of whether it is mere coincidence that left-to-right derivations track the order in which sentences are spoken and understood. It is also natural to ask how left-to-right derivations impact the psychological commitments of grammatical theories. Are they procedural descriptions of how speakers put together sentences in real time (either in comprehension or in production)? Do they amount to a retreat from linguists’ traditional agnosticism about ‘performance mechanisms’? These are questions about what a grammatical theory is a theory of, and they are the proverbial elephant in the room in discussions of left-to-right derivations in syntax, although the issues have not been explored in much detail. Here we summarize the current state of some of the evidence for left-to-right derivations in syntax, and how this relates to a number of findings by our group and others on the nature of real-time structure building mechanisms. Some of these questions have been aired in previous work (e.g., Phillips 1996, 2004), but we have come to believe that the slogan from that earlier work (“the parser is the grammar”) is misleading in a number of respects, and we offer an updated position here.

We start, however, with some important preliminaries on the question of what it is that grammatical theories aim to describe.

2. (Psychological) Goals of Grammatical Theory

Most contemporary linguists consider their theories to be theories of mental phenomena (and those that do not are of no concern to us here). What is less clear is what specific mental phenomena grammatical theories aim to describe. A standard response to this question is that grammatical theories aim to describe what a competent native speaker knows when he knows a language. This is often extended to cover the knowledge that a child brings to the language learning task, ensuring that he can arrive at the adult state. This is all very well as far as it goes. In the adult state, the knowledge in question is whatever—it-is that underlies a speaker’s ability to reliably classify sentences as acceptable or unacceptable. But this mission statement leaves us with little guidance on how to interpret the specific components of the theories that we encounter. When we are told, for example, that the wh-word what is initially merged with a verb and subsequently moved to a left peripheral position in the clause, what claim is this making about the human language system? When this is described as part of a ‘computational system’, does this mean that there is a mental system that explicitly follows this sequence of operations? In our experience this is not a question that grammarians are typically eager to discuss, but as far as we can tell the answers to the question fall into roughly three types.

First, bottom-to-top 'cyclic' derivations may be understood as literal descriptions of sequences of mental operations that speakers may, in principle, carry out. They are just unlikely to do so in the normal course of speaking and understanding. This literalist position amounts to the claim that the grammar is one among a series of structure building systems that competent speakers have, distinct from but in some fashion related to structure-building systems for comprehension and production. It is not easy to find clear endorsements of this position, although it is a perfectly clear
cognitive hypothesis. We think that this is roughly what Townsend and Bever (2001) have in mind for their dual system model of sentence parsing, which analyzes all sentences using two different syntactic systems.

A second possibility is that the sequence of steps in a grammatical derivation may be understood not as a temporal ordering of structure-building operations, but rather as a formal relation between a set of structural representations. Under this \textit{formalist} view, the full-blown representation of a sentence is a set of structural representations that are formally related to one another by the fact that they could, in principle, be sequentially ordered as a derivation. Speakers may even construct this full-blown set of representations in the course of everyday speaking and understanding, e.g., as the output of the comprehension process. But this view does not entail that the derivation describes the temporal sequence of steps that are used to assemble the full-blown representation. This is another position that we have not found frequent endorsements of in print, but it is certainly one that we have encountered in our discussions.

The third possibility is that a grammar is merely an abstract characterization of a function whose extension is all and only the well-formed sentences of a given language. Under this \textit{extensionalist} interpretation of grammatical theory, a bottom-to-top grammatical derivation is not a hypothesis about a sequence of representations that a speaker would ever mentally construct, on any time scale. Consequently, theories that adopt such derivations should not be evaluated based on their correspondence with the mental computations that speakers actually carry out. Instead, such theories are accountable only to how successfully they partition the grammatical and ungrammatical sentences of a language. We suspect that this is a position that many practicing syntacticians are comfortable with, and it certainly corresponds with much standard practice in linguistics.

The literalist and formalist positions are clear mentalistic hypotheses about grammatical derivations, and so they are amenable to empirical scrutiny based on what speakers actually do. We will have more to say below about what speakers do. On the other hand, the extensionalist position is more elusive. By claiming that it is merely an abstract characterization of a function that generates the grammatical sentences of a language, it places itself beyond the reach of most empirical evidence, aside from acceptability judgments. If one takes this position seriously, then the individual components of a grammatical theory should be understood as having no independent status as mental objects or processes, as they are merely components of an abstract function, rather than components of a more concrete description of a mental system. Despite this, we encounter many extensionalist theories nowadays that appeal to notions of "efficiency" and "computational economy" of derivations, in ways that are hard to reconcile with the notion of an abstract functional description. In fact, the extensionalist position, when taken literally, undermines even the simple notion that children learn the individual parts of their target grammar, since that would amount to treating the components of the functional description as if they are actual mental objects. Of course, one is always free to avoid these consequences by moving towards a literalist or formalist position, accepting the additional mentalistic commitments that those positions entail.

Of greatest relevance to our current concerns is the relation between the goals of extensionalist grammatical theories and the putative attractions of left-to-right structure building mechanisms. If the goal is simply to describe an abstract function that generates all of the grammatical sentences of a language, with no regard to how speakers might actually generate sentences in real time, then it may be considered irrelevant to compare the grammar's derivations to the operations of comprehension or production systems. One could even argue that the comparison is misleading, since it amounts to a comparison of theories at two different levels of description, like comparing mathematical theorems with a pocket calculator. Under this view, it would be inappropriate to seek grammatical theories whose derivations resemble actual
mental computations, since that would go beyond what a grammatical theory is supposed to do: grammatical theories with bottom-to-top or left-to-right derivations, or with no derivations at all, are accountable only to their success in partitioning the good and bad sentences of a language. In the well-known terms of Marr (1982), the extensionalist position is that a grammatical theory corresponds to a description of the language system at the computational level (as opposed to lower level algorithmic and implementational descriptions).

In practice, the relevance of real-time processes to extensionalist grammatical theories varies, because linguists have different motivations for seeking abstract functional descriptions. For what we will call a **strategic extensionalist** the description of an abstract system that generates all of the sentences of a language is not an end in itself, but it is a reasonable interim goal, and one that can be pursued relatively easily and efficiently for a wide range of languages. For the strategic extensionalist, the aim is to ultimately move beyond the functional description to a more detailed, mechanistic understanding of the human language system. Consequently, the strategic extensionalist has every reason to be interested in constructing a theory that lends itself to progressing towards this next step. On the other hand, for a **principled extensionalist** the extensionalist enterprise is an end in itself, which remains relevant even if others are able to provide lower-level characterizations of the human language system. For the principled extensionalist it is important to distinguish what the human language system computes – this is viewed as the linguist's task – from how speakers carry that out in practical situations, which is something for psycholinguists to figure out. For the principled extensionalist, it is interesting if a left-to-right derivation happens to yield better coverage of acceptability judgments, but any resemblance between such derivations and real-time processes is irrelevant. In our view, the appropriateness of the strategic or the principled extensionalist approach depends on the question of whether the human language system is implementation (in)dependent.\(^1\)

3. Implementation (In)dependence

In describing a biological or artificial system it is often useful to describe what the system does in abstract terms, and this is precisely what generative grammars aim to do. When a digital computer carries out a simple arithmetical calculation, it is certainly useful to be able to describe the calculation in abstract terms (e.g., \(48 \div 8 = 6\)) without reference to the algorithm that the computer uses to solve division problems or to the electronics of the chip that implements that algorithm. Similarly, it is useful to describe the contrast between English and Chinese wh-question formation in abstract terms (e.g., "English wh-phrases appear at the left edge of the interrogative clause, whereas Chinese is a wh-in-situ language"), setting aside questions of how English and Chinese speakers actually go about constructing wh-questions in real time, and ignoring the details of the neural processes associated with question formation.

We certainly do not dispute the usefulness of abstract descriptions, but it is important to distinguish between abstractions that are implementation independent and abstractions that are implementation dependent. An abstraction is **implementation independent** if the exact same abstract system can be realized in different ways by multiple lower-level implementations, with no change in the abstract system itself. Mathematical systems are straightforward examples of this: simple arithmetical calculations may be implemented using a variety of equivalent algorithms and using a host of different types of hardware (digital computer, abacus, human brain, etc.), with no impact at all on the nature of the arithmetical facts. In contrast, an abstraction is

\(^1\) Jackendoff (2002, ch. 2) offers a useful discussion of these issues with an added historical perspective. Jackendoff distinguishes between ‘soft’ and ‘hard’ idealizations.
implementation dependent if it is only ever realized in one way at a less abstract level of description.

Abstract generative grammars have traditionally been studied under the assumption that they are implementation independent. It is assumed, for example, that eventually we will be able to write software or build hardware to implement the human language system in computers that differ from human ‘hardware’ in many respects. As such, this is taken as evidence for implementation independence. This, however, is clearly an empirical question. It is a real and attractive possibility that computational modelers of the future might succeed in that effort, but it is far from having been accomplished at this time. The fact that some grammars that approximate fragments of human language have been digitally implemented does not settle the question. A relevant model should minimally be completely descriptively adequate, and one might reasonably want the model to approximate human linguistic abilities in more ways than simply classifying (un)acceptable sentences. Such models currently inhabit the realm of thought experiments. Here we disagree with Neeleman & van de Koot (2010), who agree with us on the importance of implementation (in)dependence, but take it as quite significant that some natural language grammars could be implemented in different artificial devices, concluding that human language is implementation independent.2

In the meantime, we argue that the emphasis on the implementation independence of generative grammars is misplaced if the purpose of the endeavor is to understand the nature of the human system. If a research community makes the choice to carve out one sub-area of the study of human language abilities, e.g., classifying (un)acceptable sentences, and then finds that this sub-area can be described in implementation independent terms, then this surely does not entail that the sub-area corresponds to a privileged, implementation independent sub-area of the human language faculty. There is no doubt that it is interesting and useful to develop explicit computational models of human language, but such models cannot show whether human grammatical abilities are, in fact, independent of their cognitive or neural implementation. Regardless of what one’s personal hunch is about the likelihood of successful non-human implementation of language, a more interesting question for our purposes is whether it is implementation independent within humans.

According to the standard formulation of the principled extensionalist view, speakers of a language have knowledge of the (un)acceptable sentences of their language, and this knowledge can be described at an abstract level by a generative grammar. Importantly, it is assumed that this abstract knowledge is implementation independent, and that speakers can put their knowledge to use in different ways in activities such as speaking, understanding, and internally generated speech. This means that for any given well-formed sentence structure defined by the grammar a speaker may have a number of different ways of mentally assembling the structure. If this assumption is supported, then the principled extensionalist is justified in separating speakers' knowledge of what is well formed from their knowledge of how to assemble well-formed structures. But if speakers do not, in fact, have multiple ways of constructing

2 Neeleman and van de Koot describe a thought experiment in which a pair of computers with radically different underlying architectures are both able to pass a kind of Turing Test for human language, such that their performance is indistinguishable from human native speakers of the language. They argue that one would want to say that both computers speak the relevant language, and conclude that this is because the computers and the humans would share an abstract, implementation-independent grammar. We agree with the intuition behind this thought experiment, but it seems to reveal more about the common usage of predicates like “speaks French” than about the nature of the human language faculty. Despite this misgiving, we find much to like about Neeleman and van de Koot’s discussion of the challenges involved in understanding grammatical theories at different levels of description.
the same representations, then human language appears to be more implementation dependent.

The evidence on whether speakers have multiple ways of constructing the same sentence representation is not extensive, but there are a number of reasons to think that speakers have one and only one way of assembling any individual representation. There is almost no evidence for the alternative view that speakers have multiple ways of building the same representation. First, speaking and understanding proceed in the same (roughly) left-to-right fashion. Although they have different goals, they have a great deal in common, appear to construct the same representations, and plausibly do so in the same order, although this topic has not been investigated in great detail. Second, in comprehension and production there is much evidence that speakers build structures and interpretations incrementally, in roughly the order in which words are presented. We are unaware of evidence that speakers are able to construct the same interpretation in different orders. For example, reading backwards is a task that lies somewhere in the difficult-to-impossible range, despite its correspondence with the derivational order assumed in many generative grammars. Third, in cases of reanalysis in sentence understanding, where comprehenders are led into a ‘garden path’ and must then reorganize their initial parse, some evidence suggests that speakers repair by returning to the error point and simply re-parsing the sentence in exactly the same order that it was presented (Inoue & Fodor, 1995; Grodner et al., 2003).

We therefore adopt the working hypothesis that natural language grammars are implementation dependent with respect to how sentences are assembled: there is only one algorithm for structure-building across all human speakers (of the same language). We would certainly welcome more systematic evidence, but currently the evidence for the alternative implementation independent position is practically non-existent. The implementation dependent position is the simpler and more falsifiable hypothesis, and hence should be preferred until proven otherwise. Consequently, we think that the principled extensionalist position is unwarranted, and that the motivations for developing abstract generative grammars are more pragmatic than principled. We have no problems with pragmatic motivations, and we recognize the value of focusing on characterizing good/bad sentences as a way of making headway in describing human language. But this is very different from the position that the characterization of good/bad sentences is a fundamentally separate enterprise from understanding real-time language processes. Ultimately we seek theories that capture how sentences are put together, and not just what their final form is.  

In discussions of these issues we sometimes encounter an objection in the form of a 'slippery slope' argument. If human language is better described in terms of real time cognitive processes rather than abstract functions, so the argument runs, then why stop there – why not continue all the way down to the level of neurochemical processes in brain cells or beyond? As with other slippery slope arguments, the expectation is that we should find this consequence appalling, and hence should drop the entire argument. We acknowledge the concern, but disagree with this argument. First, we are not arguing against the usefulness of abstract descriptions in the study of language. We find them exceedingly useful. We are simply arguing that there is no privileged level of abstraction – the level occupied by most current generative grammars – that is

---

3 Although we have focused our attention here on the interpretation of derivational grammatical theories, the issue of implementation (in)dependence is equally relevant to interpreting grammatical theories that assume no derivations. Non-derivational syntactic theories, which are often presented as preferable due to their order-neutrality, imply implementation independence. If syntactic representations are not, in fact, implementation independent, then that should count against non-derivational theories, just as it counts against bottom-to-top derivations.
exempt from considerations of psychological implementation. Second, it should be emphasized that implementation (in)dependence is not an all-or-nothing property of an abstract system. Rather, a system’s implementation (in)dependence must be evaluated at each successive degree of abstraction. Our focus in the current discussion is on the relation between symbolic descriptions of the structure of sentences and symbolic descriptions of the procedures for assembling sentences in real time. This leaves open the possibility that human language is implementation independent at a lower level. For example, there are interesting arguments that the basic notion of hierarchical structure in language must be implementation independent with respect to its neural encoding, because of the very different demands of immediate vs. long-term encoding of sentences in memory (Jackendoff, 2002; van der Velde & de Kamps, 2006). Briefly, the most plausible method for long-term information storage in the brain — through changes in synaptic connectivity — is too slow for the millisecond-scale processes needed for real-time language use. The need for dual encodings of the same structures suggests that structured representations are not always neurally encoded in the same fashion, and hence that they are implementation independent. It is far from certain that this argument goes through, because it is not clear that immediate and long-term encodings of sentences in memory are genuinely isomorphic to one another. But this is the type of argument that might one day lead to a clear finding of implementation independence for human language (for one specific level of implementational detail). In sum, the slippery slope argument is misplaced in the current case.

We should also emphasize that we do not claim that a grammatical theory is inherently preferable if it can be transparently linked to real time processes. Standard considerations of descriptive adequacy (i.e., generating all and only the acceptable sentences of a language) are as relevant as ever, and all theories should be accountable to that standard. A grammatical theory that achieves impressive descriptive adequacy using mechanisms that are opaque to real time implementation is an interesting theory nonetheless. However, the descriptive success of such a theory begs for further analysis of why it works so well, and whether its success crucially depends on the opaque mechanisms. This analysis should ideally lead to development of a more psychologically transparent version of the theory. We take up this challenge below for arguments that have been presented in the syntax literature in favor of bottom-to-top derivations.

4. Psychological Aims of a Real-time Grammar

Having examined the psychological commitments of standard generative grammars, it is appropriate to apply similar scrutiny to grammars that adopt roughly left-to-right derivations and that aspire to be models of real time processes. Here it is important to address some possible misconceptions. (One of us bears some blame for spreading some of the misconceptions.)

First, the slogan ‘the parser is the grammar’ (Phillips, 1996) sounds nice enough, but it is unfortunately misleading, as it too closely identifies the grammar with the task of language comprehension. It would probably be more appropriate to regard a real-time grammar (the ‘structure builder’) as one important component of a parsing/comprehension system, but certainly not the whole system. We envision a structure building system that combines words and phrases to form sentences and meanings in real-time, in essentially the same manner in comprehension and production. This system is task-neutral, and it could even operate in the absence of external input or a communicative goal, i.e., neither comprehension nor production is necessary. The system is, of course, put to use in speaking and understanding, where the structure that it builds is constrained by the external input (comprehension) or by the speaker’s message (production), but these tasks require far more than just a structure building system. Parsing requires some way of using incoming sounds and words to determine which specific structure building operations are appropriate.
Production requires some way of mapping the elements of an intended message to operations in the structure building system. It appears that the human parser is fairly good at mapping external input onto appropriate structure building operations, and the human production system appears to be similarly effective at selecting structure building operations that are appropriate to the speaker's communicative goals. But the assumption that natural language grammar has the form of a real-time structure building system does not logically entail the success of either parsing or production. Consequently, some caution is needed when using evidence from parsing or production to assess the nature of the structure builder.

Second, and closely related to the first point, the claim that the mental grammar should be understood as a real-time structure generation device does not guarantee that it is part of a perfect parsing device. When we claim that the real-time structure-building device is the mental grammar, we predict that the representations that this device constructs should be grammatically well formed, and that it should incorporate whatever detailed grammatical machinery we would normally expect of a grammar. This means that the system should not construct the rough-and-ready representations that some have argued to be created in the service of rapid, efficient comprehension (e.g., Ferreira & Patson, 2007; Townsend & Bever, 2001). However, the claim that the representations that are built on-line are grammatically precise does not entail that they are the same ones intended by the speaker. In extreme cases where a listener is distracted or in a noisy environment, he might use his mental grammar to construct a perfectly reasonable representation that bears only a weak relation to the sentence that the speaker uttered. This would show that the listener is failing to make full use of the input, but it would not license any clear conclusions about the nature of the listener’s real-time structure building system. Consequently, when we examine studies of on-line satisfaction of grammatical constraints, the key prediction for current purposes is that the comprehender constructs grammatically well-formed representations, even if those representations are not a grammatically possible parse of the incoming sentence. We will have more to say on this point in a moment.

Third, the claim that the grammar has the form of a real-time structure building system is independent of long-standing psycholinguistic questions about how speakers resolve syntactic ambiguities in language comprehension. Structural ambiguities arise when the input to the comprehension system has two or more well formed structural analyses, i.e., they are cases where the grammar alone cannot decide. Ambiguity resolution has enjoyed a prominent position in psycholinguistic research on sentence comprehension, and there have been long and heated debates over which types of information are brought to bear in resolving ambiguity (e.g., Frazier, 1987; MacDonald, Pearlmutter, & Seidenberg, 1994; van Gompel & Pickering, 2007). There is, in fact, broad consensus that simpler parses are favored in ambiguity resolution. The controversy surrounds what it means to be ‘simpler’ (e.g., structurally simpler, more frequent, more semantically or pragmatically natural, etc.) and what it means to be ‘favored’ (i.e., the unique parse vs. highest ranked parse among multiple parses pursued in parallel). These discussions are interesting, but they are orthogonal to our claims about the form of the grammar.

Fourth, we have been surprised at how often claims about the procedural nature of the grammar are interpreted as claims that grammatical phenomena are epiphenomenal. In linguistics and psycholinguistics one often encounters proposals that some phenomenon that has traditionally been analyzed in formal terms can instead be explained in terms of constraints on parsing or production (e.g., Kluender & Kutas, 1993; Hawkins, 1994; Harris & Bates, 2003; Hofmeister & Sag, 2010). These reductionist accounts of linguistic phenomena are often referred to as ‘processing accounts’, as a way of contrasting them with ‘formal’ or ‘grammatical’ accounts. They are claims that the phenomenon in question does not fall under the purview of the mental grammar. Reductionist analyses of grammatical phenomena are interesting, and we find some types of evidence more compelling than others (for discussion see
Sprouse, Wagers, & Phillips, 2012; Phillips, 2013ab), but they are quite different than the models that we are advocating here. Since reductionist accounts are explicitly non-grammatical in nature, they clearly have little in common with the claim that grammatical derivations follow an order that happens to be well suited to comprehension and production.

Finally, we should clarify the reason why we repeatedly describe grammatical derivations as proceeding in a *roughly* left-to-right order. Sentences are spoken and heard in a strict left-to-right order. (Strictly speaking this is a tautology, as left-to-right order is merely a conventional representation of the temporal order of words in speech.) However, it is probably not the case that mental structure building operations perfectly follow the linear order of a sentence, whether in comprehension or production. To take just one example, in a head-final language such as Japanese it may be necessary for the structure building system to create a position for the head of a phrase before it has completed the arguments and adjuncts that precede the head. More generally, structure building in comprehension is probably not entirely synchronized with the appearance of words in the input. There is growing evidence that comprehenders often build structural positions in their parses before encountering the words in the input that phonologically realize those positions (Aoshima, Phillips, & Weinberg, 2004; de Long, Urbach, & Kutas, 2005; Lau et al., 2006; Mazuka & Itoh, 1995; for review see Lau, 2009), and some evidence for related effects in production (Momma, Slevc, & Phillips, 2013). The upshot of this is that it may not even be desirable to insist upon a strict left-to-right order for grammatical derivations, since the operations of the real-time structure builder may not proceed in a strict left-to-right order. If it is the case that there is a single structure-building system that assembles sentences in a strict order, then it is likely that this order will turn out to be only *roughly* left-to-right. What matters is whether the order of structure building operations is consistent.

5. **Is On-Line Structure Building Grammatically Precise?**

If the representations that are built during real-time comprehension and production differ from those that are motivated by standard grammatical analysis, then we face *prima facie* evidence for multiple structure building systems, and thus evidence against our claim of a single procedural grammar. Fortunately, there is now much experimental evidence that bears on this question, and it mostly supports the notion that real-time processes assemble syntactic representations that are the same as those motivated by grammatical analysis. We should reiterate here that the question of how grammatically sophisticated real-time structure building is must be kept distinct from the question of how accurately the parser analyzes the input. A parsing system could, in principle, build wonderfully rich grammatical representations that are nevertheless rather poor analyses of the incoming strings of words.

Many classic psycholinguistic studies from the 1960s and 1970s addressed the question of whether on-line language processes create representations with the detailed structural properties that were proposed in the newly emerging field of generative grammar. This body of work is most often remembered for its failure to find correlates of the transformational derivations found in theories of that time, but a

---

4 The actual findings from these studies were less definitive than is typically claimed in historical reports (e.g., Townsend & Bever, 2001, ch. 2). The tests focused on a rather narrow linking hypothesis, and many of the conclusions relied on specific grammatical analyses that have not stood the test of time. Tests of the 'Derivational Theory of Complexity' (DTC) examined the narrow linking hypothesis that the perceptual complexity of a sentence is best predicted by the number of steps in its transformational derivation. But given the many different factors that contribute to the perceptual complexity of a sentence, there is little reason to expect that derivation length should be the primary
more enduring outcome is that the same studies often found experimental support for
the surface structure representations of the time (for reviews see Fillenbaum, 1971;
Fodor, Bever, & Garrett, 1974; Levelt, 1974). For example, hierarchical clustering
analyses of speakers’ relatedness judgments for word triads from a sentence yielded a
good approximation to the surface structure of a sentence (Levelt, 1970), and studies
on the perceptual mis-location of non-speech clicks played during a sentence
confirmed a difference between object control sentences like The general defied the
troops to fight and exceptional case marking constructions like The general desired
the troops to fight (Bever, Lackner, & Kirk, 1969).

Far more recently, studies that use highly time-sensitive measures such as event-
related brain potentials (ERPs) have made it possible to track how quickly
comprehenders are able to detect different types of anomaly in the linguistic input.
This work has shown that speakers detect just about any linguistic anomaly within a
few hundred milliseconds of the anomaly appearing in the input. Different types of
grammatical anomalies elicit one or more from among a family of different ERP
components, including an (early) left anterior negativity (‘(e)LAN’; Neville et al.,
1991; Friederici, Pfeifer, & Hahne, 1993) or a P600 (Osterhout & Holcomb, 1992;
Hagoort, Brown, & Groothuizen, 1993). Many questions remain about what the
different components reflect and what determines which components are evoked in
any individual situation (Hagoort, 2003; Friederici & Weissenborn, 2007; Lau,
Phillips, & Poepppel, 2008; Federmeier & Laszlo, 2009; Gouveia et al., 2010; Brouwer,
Fitz, & Hoeks, 2012), but for current purposes the most relevant outcome from this
research is that more or less any grammatical anomaly elicits an ERP response within
a few hundred milliseconds. If the on-line analyzer is able to immediately detect any
grammatical anomaly that it encounters, then it is reasonable to assume that it is
constructing representations that include sufficient grammatical detail to detect those
anomalies.

Another body of on-line studies has examined whether on-line structure building
respects various grammatical constraints, i.e., whether the parser ever creates
grammatically illicit structures or interpretations. Many studies have found evidence
of immediate on-line effects of grammatical constraints, such as locality constraints on
wh-movement (Stowe, 1986; Traxler & Pickering, 1996; Wagers & Phillips, 2009),
and structural constraints on forwards and backwards anaphora (Kazanina et al., 2007;
Nicol & Swinney, 1989; Sturt, 2003; Xiang, Dillon, & Phillips, 2009; Lewis, Chow, &
Phillips, 2012; Dillon et al., in press). These findings extend to complex cases that
present apparent challenges for incremental application of grammatical constraints,
such as constraints on backwards anaphora in Japanese, where the constraints must
apply before any verb has appeared in the input (Aoshima, Yoshida, & Phillips, 2009),
and constraints on parasitic gaps inside complex subjects in English, where the
parasitic gap precedes its licensor (Phillips, 2006). Findings such as these imply that
the structures created on-line include sufficient structural detail to allow the
constraints to be applied during parsing.

The many different types of evidence for on-line grammatical sensitivity do not, of
course, strictly require that the structures that are built on-line are exactly those that

are sanctioned by the grammar, or that the system that builds them is a procedural grammar. It is always possible that the on-line structure builder is not, in fact, identical to the grammar, but instead is a very effective ‘covering grammar’ for the true grammar of the language. But to the extent that the on-line structure builder constructs exactly the right representations to capture both on-line behavior and standard acceptability judgments, then we see little motivation for postulating an independent grammar that yields no additional empirical coverage.

There are, however, a number of findings in the psycholinguistics literature that have been taken to indicate divergence between the structures created on-line and those motivated by traditional grammatical analysis. We will briefly review three types of evidence, from misinterpretations, syntactic priming, and grammatical illusions.

Comprehenders frequently misinterpret the sentences that they encounter. Fernanda Ferreira and colleagues have argued that this is a desirable property for a comprehension system, and that the misinterpretations are the result of ‘good enough’ (GE) representations, which contrast with representations that are “detailed, complete, and accurate with respect to the input” (Ferreira & Patson, 2007, p. 71; see also Ferreira, Bailey, & Ferraro, 2002). One recent event nicely illustrates the intuition behind Ferreira’s argument. On October 2nd 2008 a record TV audience (70 million in the US alone) watched the debate between the vice presidential candidates, Senator Joseph Biden of Delaware and Governor Sarah Palin of Alaska. In a segment on energy and climate change Palin said, “I’m not one to attribute every activity of man to the changes in the climate”. Probably most viewers immediately recognized what Palin intended to convey with this quote (that climate change is not primarily caused by human activity), and few would have even noticed that the literal interpretation of her statement is quite different (that human activities are not exclusively caused by climate change). Palin’s slip-of-the-tongue was (roughly) a classic exchange error, and comprehenders successfully recovered Palin’s message despite the error, presumably because they had a good idea of what the message was likely to be. Ferreira and colleagues argue that situations like this show GE representations in action, and straightforwardly illustrate one of the benefits of an interpretive system that is not a slave to the precise surface form of incoming sentences.

Ferreira and colleagues offer additional evidence of systematic misinterpretations from experimental studies. For example, they tested what interpretations speakers take away from garden path sentences in which they initially misparse and must then reanalyze, e.g., While Anna dressed the baby played in the crib (Christianson, Hollingworth, Halliwell, & Ferreira, 2001). When questioned after the sentence, speakers reliably agreed that the sentence stated that the baby had played, but they also agreed on a substantial proportion of trials that the baby was dressed, suggesting that in their final interpretation the single NP the baby filled two conflicting thematic roles. They refer to this as a ‘lingering’ garden path effect (see also Sturt, 2007). Similarly, Ferreira reports that comprehenders often incorrectly judge that the surface subject is the ‘do-er’ in passive sentences like The dog was bitten by the man (Ferreira, 2003).  

We agree with Ferreira that the misinterpretation of good sentences and the successful repair of speech errors are important phenomena that show that on-line interpretation does not always deliver the correct meaning of an incoming sentence. But as we have

---

5 In this study comprehenders were probably not simply following a plausibility-based heuristic, as they also showed very similar error rates in judgments of non-reversible passives, such as The cheese was eaten by the mouse. Only in the most difficult sentence types (object clefts) did error rates increase in the manner predicted by a plausibility-based heuristic.
emphasized above, there is a difference between failure to correctly analyze the incoming string of words and failure to build a grammatically well-formed representation in response to the input. As psycholinguists we are certainly very interested in the question of how effectively people parse, but this is separate from the question of whether the on-line structure builder generates only grammatically well-formed representations. The evidence reviewed by Ferreira does not show that comprehenders assemble grammatically ill-formed representations. We suggest that in the case of ‘lingering interpretations’ of garden path sentences what happens is that comprehenders incrementally update their interpretations over the course of the sentence, but that interpretations are not labeled based on the pieces of syntax that generated them. Consequently, when a syntactic parse undergoes reanalysis to a complete and well-formed structure, rescinding of the incorrect syntactic analysis does not lead to rescinding of any interpretations that were previously generated by that syntactic analysis. Interpretive repair of speech errors, as in the example from the vice presidential debate, presents a slightly different situation, since the repairs occur in situations where the comprehender has a good idea of the intended meaning and hence the speech input is redundant. In such cases the interpretive system likely completes its task before the syntactic parse has finished. It is interesting that the interpretive system is able to do this, but it tells us little about the nature of the syntactic representation that is generated in these situations, which may be perfectly grammatically well-formed. A related observation that may be relevant here is that different types of speech errors appear to impact language comprehension in different ways. In sentences that are syntactically appropriate but semantically garbled, comprehenders are sometimes able to recover the intended interpretation without even noticing the error, as in the example above. In contrast, when speech errors lead to syntactic deformation of a sentence, comprehenders are generally able to recover the intended form, but the error does not pass unnoticed. This suggests that comprehenders may be able to skip detailed interpretation of an incoming sentence, but that they cannot skip syntactic analysis, even if it is not needed for interpretation.

A second potential argument for a mismatch between on-line structure building and the structures sanctioned by the grammar comes from syntactic priming. In language production, many studies have shown that the use of a particular syntactic structure increases the likelihood that the same structure is used in subsequent utterances, even when there is no lexical overlap between the utterances that have the same structure (Bock, 1986; Pickering & Ferreira, 2008). This finding has led to interesting questions about what pairs of structures count as the ‘same’ for purposes of syntactic priming. An influential study by Bock and Loebell (1990) showed priming between VPs containing PPs that differ in their thematic roles, e.g., VPs with a locative PP such as *The wealthy widow drove the Mercedes to the church* primed VPs with a recipient PP such as *The wealthy widow gave the Mercedes to the church*, and VPs with a locative by-phrase such as *The 747 was landing by the control tower* even primed full passives such as *The 747 was alerted by the control tower*.(Note that we use lexically matched examples here for illustrative purposes only. The adjacent prime-target pairs in these studies did not overlap in this way.) In addition, many other studies have shown that syntactic priming does not depend on overlap in the function words (including prepositions) between primes and targets (Bock, 1989; Fox Tree & Meijer, 1999). Taken together, the evidence therefore suggests that relatively coarse-grained syntactic parallelism is sufficient to cause structural priming. This must somehow be

---

6 This account clearly begs the question of how the relevant interpretations are accessed or generated, if not through detailed compositional interpretation of the sentence structure. Investigation of this question may well lead to the conclusion that the language comprehension system has multiple ways of using cues to generate interpretations. However, this is quite different from the conclusion that there are multiple representational systems or multiple real-time structure building procedures.
reconciled with the evidence that motivates linguists to postulate fine-grained structural distinctions between superficially similar sentences. Different responses to this challenge are possible, and they remain to be resolved. The structural priming evidence could be taken as support for grammatical models that make less fine-grained structural distinctions (Culicover & Jackendoff, 2005). Or it could be taken to show that the representations involved in structural priming are not those defined by the grammar (we do not favor this option, but it is a logical possibility). Alternatively, it could be taken to show just that the structural priming paradigm is a relatively blunt tool for investigating structure, because relatively coarse-grained similarity between structures is sufficient to cause structural priming. This would leave open the possibility that on-line processes build fine-grained structures.

A third potential motivation for distinguishing the structures built on-line and those sanctioned by the grammar comes from grammatical illusions, cases where comprehenders appear to fleetingly accept structures that are judged bad after more reflection. The most notorious case of a grammatical illusion involves comparative constructions such as More people have been to Russia than I have, which are semantic gibberish but initially sound remarkably good. Townsend and Bever (2001) argue that such cases make an interesting case for a system that distinguishes a rough-and-ready initial analyzer from the more fine-grained analyses of the grammar. A number of additional cases of illusory acceptability have emerged in recent years. One case involves the spurious licensing of the negative polarity item (NPI) ever by non c-commanding negation, as in The bills that no senators have supported will ever become law (Drenhaus, Saddy, & Frisch, 2005; Xiang, Dillon, & Phillips, 2009). Another case involves illusory agreement licensing in which a plural-marked verb is judged to be acceptable in the vicinity of a plural NP that is not its syntactic subject, as in The runners who the driver see ... or The key to the cabinets probably are ... (Clifton, Frazier, & Deevy, 1999; Pearlmuter, Garnsey, & Bock, 1999; Wagers, Lau, & Phillips, 2009). Yet another case involves evidence that during the processing of pronouns comprehenders fleetingly consider a clause-mate subject NP as a potential antecedent, in violation of Binding Principle B (Badecker & Straub, 2002; Kennison, 2003), although findings are mixed, and a number of other on-line studies report immediate effects of Principle B (Clifton, Kennison, & Albrecht, 1997; Lee & Williams, 2006; Nicol & Swinney, 1989; Runner, Sussman, & Tanenhaus, 2006; Lewis et al., 2012). For a more detailed review of where grammatical illusions do and do not arise see Phillips, Wagers, & Lau (2011).

Grammatical illusions pose a challenge to our hypothesis of a real-time procedural grammar to the extent that they justify a mismatch between on-line structure building processes and processes that operate under less time pressure. However, demonstrations of illusory acceptability do not automatically show such a mismatch. Grammatical illusions could have a number of sources other than the existence of multiple structure building systems. They could reflect mis-parsing, where the comprehender builds a perfectly grammatical representation that happens to not match the input. Alternatively, they could show that the normal workings of the structure building system are slow enough that on-line methods can probe the intermediate steps of the computation. Wellwood and colleagues have recently argued that illusory comparatives reflect mis-parsing rather than construction of an ill-formed internal representation. In acceptability rating studies they find that the illusions are more robust with predicates that are ‘repeatable’ such as go to the gym than with predicates that are ‘non-repeatable’ such as won the lottery yesterday, and suggest that this is

---

7 These comparative illusions, which some have referred to as Escher sentences, were first pointed out by Montalbetti (1984), although he presents them as a curiosity and does not offer an analysis of the phenomenon, which has not been systematically studied until recently.
because the sentences are mis-interpreted as event quantification rather than as quantification over individuals (Wellwood et al., 2013). Notably, English grammar allows the syntax of individual quantification to be interpreted as event quantification in certain contexts, as in The Washington DC metro carries more than 200 million passengers per year, which is a claim about person-trips rather than about distinct individuals.

The other cases of grammatical illusions that we have listed here are probably not amenable to a mis-parsing analysis, but might nevertheless reflect the normal operations of a grammatically accurate structure building system. The temporary consideration of local antecedents for pronouns that violate Principle B (as found in some, but by no means all on-line studies of Principle B) may be directly related to a natural grammatical implementation of the constraint, applying as a filter that marks candidate referential dependencies as illicit, rather than as a constraint that prevents the generation of illicit candidates. Since pronouns may be associated with a wide range of syntactic and discourse antecedents, a mechanism that generates candidates and then excludes inappropriate candidates may be the most feasible way of applying Principle B, and it may be that the fleeting consideration of illicit antecedents found in some studies reflects this mechanism in action. More broadly, ‘generate-and-filter’ mechanisms are familiar from many grammatical theories (e.g., Chomsky, 1981; Legendre, Grimshaw, & Vikner, 2001), and hence are plausible components of a real-time grammar. The case of illusory NPI licensing remains poorly understood, as few NPIs and only a narrow range of syntactic and semantic contexts have been tested to date, so it is unclear how far illusory NPI licensing extends beyond the specific environments examined so far. However, Xiang and colleagues have attempted to show how the illusions could reflect the operations of the same pragmatic licensing mechanisms that are widely held to be responsible for normal NPI licensing (Xiang et al., 2009). According to this account, the illusions reflect inappropriate pragmatic inferences, rather than use of rough-and-ready licensing mechanisms. In the case of illusory agreement licensing Wagers and colleagues have argued that the illusions arise from the use of grammatically fully appropriate feature retrieval operations in an architecture with noisy memory representations (Wagers et al., 2009). We should emphasize that systematic investigations of grammatical illusions are relatively new and that none of these accounts are definitive. However, they show that it is possible to account for the phenomena using mostly standard structure-building and interpretation mechanisms, without recourse to independent on-line and off-line systems. One clear advantage of the single-system accounts is that they predict that only certain types of temporary illusions should be possible. Accounts that invoke a separate rough-and-ready structure building system are less constrained.

6. **Bottom-to-top derivations**

We have argued that human grammars are implementation dependent, at least with respect to how syntactic structures are assembled, and that it is desirable and plausible to view grammars as procedural systems that can be understood in terms of actual real-time mental computations. However, most contemporary syntactic theories seem incompatible with this view: they assume that grammatical derivations proceed in a strictly bottom-to-top order, progressively combining words and phrases starting from the most deeply embedded elements. For most languages this yields derivations that proceed in a mostly right-to-left order, i.e., the opposite of the order in which comprehension and production operate. We must therefore address the evidence for bottom-to-top derivations, to determine whether it undermines our approach. We have found, to our surprise, that the evidence is neither extensive nor particularly well known. The assumption of a bottom-to-top order of derivation seems to have arisen from some reasonable intuitions about human language. However, it is an assumption that has rarely been revisited over the years. Bottom-to-top is widely regarded as the only possible order of derivation, whereas it ought to be considered as just one among several possible ways to account for certain facts about language.
Probably the most influential intuition underlying this pervasive assumption is that lexical items are inserted into thematic positions. That is, a structure “begins” as a direct representation of thematic relations, which is subsequently transformed to satisfy other syntactic requirements. This intuition has its origins in the kernel sentences of *Syntactic Structures* (Chomsky, 1957), and it was codified in the Standard Theory as Deep Structure (later ‘D-structure’ or DS), a level of representation that “defines grammatical relations and functions in a straightforward manner” (Chomsky, 1970). It was the point in the derivation at which the argument selection requirements were satisfied (the ‘Theta Criterion’). Empirically, DS was crucial to the distinction between raising and control. In raising constructions like (1), the subject bears a thematic relation to the embedded verb, but not the main verb. In control constructions like (2), the subject has a thematic relation to both verbs. It was claimed that raising constructions are formed by movement from a thematic position in the embedded clause to a non-thematic position in the main clause, whereas control constructions involve an empty category in the embedded clause that is not created by movement, but rather is created by *equi-NP deletion* or, in later parlance, the empty category PRO.

(1) The man$_i$ seemed to$_i$ enjoy his ice cream.

(2) The man$_i$ tried PRO$_i$ to enjoy his ice cream.

D-Structure was argued to be unnecessary as part of the Minimalist Program: the Theta Criterion was a requirement that might as well be satisfied at LF (Chomsky, 1993). Nevertheless, the new theory essentially reinstated it at the point of lexical insertion by requiring thematic roles to be assigned only at external Merge (Chomsky, 1995; but for a different view see Hornstein, 2000).

Applying the theta criterion at the point of lexical insertion is admittedly a fine way of capturing these data, and a bottom-to-top order of derivation follows from this. In derivations of this kind, satisfaction of the thematic requirements of a lexical item is guaranteed, and all further transformational operations serve to satisfy other syntactic requirements (encoded as “features”). However, this approach is by no means required. Stripping away theory-particular terminology, the aim is to account for the fact that arguments enter into multiple syntactic relations within a single structure. Usually only one of those relations (the one lowest in the structure) is thematic, except in the case of control constructions. These facts could just as well be captured by the opposite restriction: arguments must merge into a non-thematic position that satisfies syntactic requirements such as case, agreement or scope marking, and subsequent transformations could relate that argument to one or more lower positions in the structure where thematic and other requirements are satisfied. Arguably such operations are required under either framework: even in bottom-to-top theories, the original thematic relations must be re-established at LF for interpretation.

A second intuition underlying bottom-to-top derivation is that the endocentricity of syntactic structures is established by heads “projecting” phrases. The features of a head are said to be “passed up” to the maximal projection. Again, while such terminology is a convenient way of describing endocentricity, it is certainly not required. Bare Phrase Structure theory (cf. Chomsky, 1995) highlights the possibility that different levels of projection simply encode the relations between the head (a lexical item) and other parts of the structure. In the X’-style structure in (3), it might seem necessary to ensure that the features of the head ‘travel’ up the tree to give content to the otherwise empty categorial nodes. In the bare phrase structure in (4), the maximal projection of *loves* has no need to inherit features from the lexical item: it is that lexical item, represented in relation to other lexical items. In this approach, the labeling of levels of projection is merely a mechanism for identifying the dominant node, which determines how the phrase behaves as a whole. Endocentricity is a condition on the structure as a whole, and need not be explained by requiring the
projection of features from bottom to top. Mechanisms for regulating the match between the features of heads and maximal projections are familiar from constraint-based grammatical formalisms such as Head Driven Phrase Structure Grammar (HPSG), and can be straightforwardly implemented without the need for an ordered derivation.

(3) X’ Structure

\[ \begin{array}{c}
  \text{NP} \\
  \text{N} \quad \text{V} \\
  \text{John} \quad \text{loves} \quad \text{NP} \\
  \text{loves N} \\
  \end{array} \]

(4) Bare Phrase Structure

\[ \begin{array}{c}
  \text{loves} \\
  \text{John loves} \\
  \text{loves Mary} \\
  \end{array} \]

We have argued that bottom-to-top derivation is not a logical necessity, despite the intuitive appeal of some long-standing assumptions. Still, since some version of this approach has been widely assumed for at least 40 years, we might expect that some other aspects of grammatical theory have been found to depend on it. In fact, our impression is that empirical arguments for this approach, to the extent that they are successful at all, are arguments for ordered derivations, but not arguments for a specific derivational order.

One aspect of many transformational grammatical theories that references the order of derivation is the syntactic cycle. In fact, Chomsky’s (1966) formulation of cyclicity, which he refers to as a “general convention”, is essentially a stipulation of bottom-to-top derivation. This restriction on the application of “singulary transformations” allowed the more powerful “generalized transformations” to be eliminated in favor of recursive base rules. The Strict Cycle Condition (SCC) (Chomsky, 1973) is slightly more sophisticated: it prohibits the operations in any particular syntactic domain from feeding operations in a subdomain. Some form of bottom-to-top derivational order follows from this requirement. A primary motivation for the SCC was to address Ross’ (1967) observations about syntactic “islands”, or barriers to long-distance movement. In a strictly bottom-to-top derivation, it is not possible to insert additional structure between a moved element and its original position after a licit movement is complete. That is, an element cannot escape being a barrier to a long-distance dependency simply by being added to the structure after the dependency is formed.

Freidin (1978) points out that though the empirical coverage of the SCC is substantial, it is subsumed by several other independently-motivated conditions on transformations. Freidin covers a number of different constructions; here we will mention just one example. The derivation in (5), resulting in the illicit sentence in (6), may be analyzed as a violation of the SCC.
(5)  
   a. \[CP [IP John knows [CP [IP who_i saw what_j]]]]
   
   b. \[CP [IP John knows [CP who_i [IP t_j saw what_j]]]]
   
   c. \[CP who_i [IP John knows [CP t_j [IP t_j saw what_j]]]]
   
   d. \[CP who_i [IP John knows [CP what_j [IP t_j saw t_j]]]]

(6)  *Who does John know what saw?

In this derivation, the lower wh-phrase (what) moves after the higher wh-phrase (what) has already moved to the top of the tree. The movement in (5c) of a wh-phrase to the higher CP precedes the movement in (5d) of the other wh-phrase to the lower CP. The SCC rules out this derivation, since the lower CP is a cyclic subdomain of the higher CP. However, that step of the derivation also involves the deletion of a bound trace, which would presumably result in additional problems. Freidin invokes a “Trace Erasure Prohibition,” which would rule out sentences like (6) regardless of the SCC. One could argue that Freidin’s various conditions and principles are themselves stipulative, but his overall point is well taken, that cases that are excluded by the SCC are often already handled by independent constraints.

More generally, even if we allow that the examples in (5-6) motivate an ordered derivation in order to close off the theoretical loophole that might otherwise allow derivations like (5), such cases do not motivate a specific derivational order. Any restriction that forces derivations to proceed in a consistent order should do the trick. A derivation that proceeded from top-to-bottom would close the loophole in exactly the same way as does a bottom-to-top derivation. A strictly left-to-right derivation would equally block the derivation in (5), although it would conceivably make different predictions in situations where linear order and c-command relations diverge. Furthermore, examples like (5) are relevant to considerations of derivational order only due to a theory-internal assumption. If the locality constraint that rules out sentences like (6) is a constraint on movement operations (e.g. “wh-phrases mustn’t move across other wh-phrases”), then there may be a need to block the derivation in (5). But if the locality constraint is a constraint on the representations created by movement (e.g., “a wh-phrase cannot form a dependency that crosses another wh-phrase”), then sentences like (6) will be ruled out irrespective of the order in which sentences are put together. This should come as no surprise: it is the reason why non-derivational grammatical theories are able to handle cases like (6).

Chomsky (1995) introduced the Extension Condition to capture the effects of cyclicity in a minimalist framework. According to the Extension Condition, all syntactic operations (i.e., Merge or Move) must target the root of the tree. Cyclicity follows immediately, as well as a prohibition on movement into complement positions. It has also been claimed that root Merge is “simpler” than other instances of Merge, and is thus preferable on Minimalist grounds. Along similar lines, Chomsky (2000, p. 136) argued that root-only merger is preferable as it has the effect that inserting new terminal nodes satisfies a ‘least tampering’ condition, because it does not destroy existing constituents. Regardless of the status of these claims, however, Chomsky (1995) acknowledged that the Extension Condition could only apply before Spell-Out: covert movement operations seemed to be exempt. Adjuncts such as adverbials were also exempt. Subsequent formulations of cyclicity, such as Richards’ (1999) “featural cyclicity,” attempted to remedy the stipulative nature of the Extension Condition.  

---

8 Lasnik (2006) points out that this argument depends on the assumption of standard tree notation, and that it does not go through if other notational conventions are adopted.
However, they did not remedy the reliance of covert processes on non-root merger. If it is not the case that all syntactic operations can rely on root Merge alone, it seems arbitrary to require that pre-Spell-Out processes are restricted to it on the basis of its supposed simplicity. Moreover, notions of what counts as ‘simpler’ are subjective and perhaps not the most effective guide to figuring out the core properties of human language.

In the context of evidence for cyclic bottom-to-top derivations one often encounters discussion of successive cyclic movement. Successive cyclic movements are operations that displace a constituent across a long distance via a sequence of shorter steps. Successive cyclic movement was proposed as a way of reconciling evidence that on the one hand shows that phrases can be displaced across an arbitrarily long distance with evidence that on the other hand shows that many relatively local displacements are illicit. The proposal was that all displacement is subject to strict locality restrictions, and that apparent long-distance displacement is the result of a sequence of local operations (Chomsky, 1973). Some interesting cross-language phenomena have been offered as phonologically overt evidence of successive cyclic movement (e.g., Torrego, 1984; McCloskey, 1991). However, the properties of successive cyclic movement are compatible with a number of different derivational (or non-derivational) schemes. Successive cyclicity has attracted new interest in recent years in the context of minimalist notions of phase-based derivations (Chomsky, 2000; Boeckx, 2008). But again, we find little in these discussions that motivates a specific derivational order. Derivations that move elements from one phase to the next can be readily translated from a bottom-to-top order into other orders.

To summarize, we have argued that the evidence for strict bottom-to-top derivations is less than overwhelming. Some motivations for bottom-to-top derivations are the result of historical accidents, some reflect aesthetic choices about ‘simplicity’ that are likely to persuade only those who are already persuaded, and others may indeed motivate strictly ordering of syntactic operations, but do not entail that the specific order be bottom-to-top. Therefore we may pursue the idea that grammatical derivations reflect real-time processes with no concern that we are ignoring strong evidence for bottom-to-top derivations.

7. Current status of arguments for (roughly) left-to-right syntactic derivations

Having spent much time discussing real-time structure building processes, and assessing the evidence for bottom-to-top syntactic derivations that look quite unlike those processes, we now turn our attention to the evidence for syntactic derivations that look somewhat more compatible with real-time mental computations.

Phillips (1996, 2003) argued that the constituency diagnostics that are the bread-and-butter of syntactic argumentation are better understood if we assume that sentence structures are assembled in a syntactic derivation that proceeds from left to right. This proposal was motivated entirely by the standard linguists’ concern for distinguishing acceptable and unacceptable sentences. The possible psycholinguistic consequences were immediately apparent, however, and the proposal spawned an extensive experimental research program that has explored the nature of real time grammatical computation. We will not reiterate Phillips’ arguments here in great detail, as the originals are readily available, but we will give an outline of the key proposal, in order to provide context for the responses that we discuss.

Phillips was concerned with a notorious problem in syntax. The syntactician’s toolkit includes a variety of different diagnostics that can be used to probe the constituent structure of sentences. These diagnostics are based upon such processes as coordination, movement, ellipsis, substitution, scope, and binding, and they are a standard part of any introductory syntax course. In certain parade cases the diagnostics converge on the same result. For example, the diagnostics provide ample evidence for
the existence of a VP constituent: VPs can be coordinated, moved, elided, or substituted, and scope and binding relations show a clear subject-predicate asymmetry. However, a fact that is rarely discussed but well known to any practicing syntactician is that in many cases the diagnostics do not converge on the same results. Some diagnostics appear to apply to a broader range of constituents than others, some yield results that seem to be at odds with the results of others, and some even appear to contradict themselves by diagnosing conflicting constituent structures.

Coordination is often used as a diagnostic of constituenthood in introductory texts, but it is used much less often in syntactic research due to the fact that it is very liberal. In addition to stereotypical VPs (7a), coordination can target the two objects of double object constructions (7b), subject-verb combinations (7c), and sequences that span two clauses (7d).

(7)  
  a. Wallace [visited Wendolene] and [bought some wool].
  b. Wallace gave [Gromit a biscuit] and [Shawn some cheese] for breakfast.
  c. [Wallace designed] and [Gromit built] an enormous tin moon-rocket.
  d. Alice [knew Fred wanted to talk] and [hoped that he wanted to argue] with the president.

Other diagnostics are less liberal. For example, many strings that can be coordinated cannot be moved. The two objects of a double object construction can be coordinated, but they cannot be moved leftward or rightward in English (8). Similarly, ellipsis is more restrictive than coordination. Ellipsis can delete material from more than one clause (9a), but it cannot target a main clause plus a subpart of an embedded clause (9b), although the same sequence can be coordinated (5d).

(8)  
  a. *[Gromit a biscuit] Wallace gave ___ for breakfast.
  b. *Wallace gave ___ at breakfast time [his favorite pet beagle an enormous chewy biscuit.]

(9)  
  a. Alice [knew that Fred wanted to talk with the queen] and Ethel did ___ too.
  b. *Alice [knew that Fred wanted to talk with the queen] and Ethel did ___ with the president.

A pair of cases that have attracted a good deal of attention involve the relation between different structural diagnostics when they apply in the same sentence. The first case is a conflict noticed by Pesetsky (1995). The VP-fronting in (10), which places a verb and its two objects at the front of a clause and strands an adverbial PP, suggests a VP structure in which the predicate give books to them is a constituent and the adverbial PP attaches above this constituent. However, under the common assumption that binding is a diagnostic of c-command relations, then the successful licensing of the reciprocal each other suggests that the adverbial PP occupies a lower position where it is c-commanded by the pronoun them. These two structural diagnostics are thus in conflict with one another, and examples like (10) have come to be known as Pesetsky's paradox. A complementary case pointed out by Phillips (1996) involves VP-ellipsis, which targets very similar strings to VP-fronting, yet appears to avoid conflicts with c-command diagnostics involving binding and scope. For example, a pronoun inside an elided partial VP cannot bind a reciprocal inside a stranded adverbial, consistent with a structure in which the adverbial is structurally higher than the elided phrase (11). We will have more to say on this contrast below.
Derivational order in syntax

John wanted to give books to them, and [give books to them] he did ___ on each other's birthdays. [VP-fronting]

*John wanted to give books to them at Christmas, and Bill did ___ on each other's birthdays. [VP-ellipsis]

The general problem of constituency conflicts has rarely been addressed directly. It has probably received the most attention in the Categorial Grammar (CG) literature, where the broad range of constituents identified by some diagnostics – coordination in particular -- has been offered as support for the CG mechanisms that allow words to be combined in different orders, leading to effects of flexible constituency (Steedman, 1985; Dowty, 1988; Pickering & Barry, 1993). Steedman (1997, 2000) offers a partial typology of structural diagnostics that accounts for why some tests yield different results than others. He argues that processes such as coordination, ellipsis, and unbounded dependencies are sensitive to the order of combinatorial operations in surface structure, and hence allow for effects of flexible constituency, whereas binding relations reflect prominence hierarchies in an independent semantic/argument structure representation that is insensitive to flexible combinatorial rules. Furthermore, Steedman’s typology is not arbitrary, in the respect that processes that impact word order and surface form are associated with surface structure, whereas processes that impact interpretation are associated with the semantic structure representation. Although Phillips (2003) criticizes Steedman’s approach on a number of grounds, arguing that it fails to predict which processes are (in)dependent of one another, the overall approach is highly appealing.

Pesetsky (1995) offers an account of constituency conflicts framed within a transformational grammar approach. Pesetsky proposes that different diagnostics yield contrasting results because sentences have more than one structural representation or derivation. He argues that every VP has a cascade syntax representation, which is largely right-branching, and a layered syntax representation, which is largely left-branching. Different structural diagnostics are assumed to be sensitive to one or the other of these representations, leading to an alternative typology of structural diagnostics. Setting aside specific empirical concerns about this approach (for discussion see Phillips, 2003), the main question that this approach raises involves the apparent arbitrariness of the typology.

Phillips (1996, 2003) offered a different approach to the typologies offered by Steedman and Pesetsky in an attempt to avoid the need to stipulate that individual structural diagnostics are sensitive to one type of representation and not another. He argued that the contrasting results of different structural diagnostics are natural consequences of the dynamics of left-to-right structure building. Starting from the assumption that humans have some kind of left-to-right structure building system for language, and assuming further that this system does not countenance vacuous projection of non-branching nodes,9 some interesting predictions emerge about the evolution of constituent structure as a sentence is assembled. Whenever a word or phrase is merged with the current structure at any point other than the root node, some constituents cease to be constituents. For example, if the sentence *Harry met Sally is assembled from left to right, then there is a stage in the derivation where *Harry met is a constituent, consisting of the entire structure, but addition of the direct object Sally

---

9 This assumption is not in conflict with experimental findings that suggest predictive creation of syntactic nodes before their phonological realization is encountered in the input. Predictive structure building in language comprehension demonstrates that structure-assembly is not fully synchronized with the words in the input. This is fully compatible with the assumption that all non-terminal nodes in syntax are branching nodes.
creates a new constituent *met Sally* while destroying the constituent *Harry met*. This means that the sequences of words that are constituents change over the course of a derivation. Phillips argues that the varying results of different constituency diagnostics are a direct consequence of this dynamic property of constituent structure, and he proposes that the results of individual diagnostics should be predictable, based on which stages of a left-to-right derivation they apply to. If a sequence of words is a constituent that does not survive to the end of the sentence, then it should only be visible to constituency diagnostics that can be applied before the constituent is destroyed.

Phillips presents a number of generalizations about structural diagnostics that he attributes to the dynamics of left-to-right derivations. For example, the liberal nature of coordination is argued to be a consequence of the fact that it places constituents adjacent to one another, and hence applies before constituents are destroyed. In contrast, ellipsis is a more restrictive diagnostic, because it involves structures in which an ellipsis site in a second clause (typically) is dependent on an antecedent in an earlier clause, and hence it can only ‘see’ the constituents that have survived to the end of the earlier clause. Meanwhile, structural diagnostics that test for c-command relations are predicted to not conflict with one another, since left-to-right expansion of a structure adds new c-command relations without destroying existing c-command relations.

Phillips also offers an account of why VP-fronting gives rise to Pesetsky’s paradox, whereas VP-ellipsis does not. In cases of VP-fronting like (10) the movement operation applies before the stranded adverbial is added to the structure, and it is only the addition of the stranded adverbial (low inside the ‘reconstructed’ VP) that destroys the constituency of the fronted constituent. Therefore, tests of movement and binding appear to conflict with one another, because they apply at different stages in the assembly of the sentence. In cases of VP-ellipsis like (9), on the other hand, no corresponding conflict arises, because at the point when ellipsis must be licensed the right-branching VP-structure required to license the binding relation in (11) no longer has the sequence *give books to them* as a constituent. That sequence is certainly a constituent at some point during the construction of the first clause, but by the time the first clause is completed it is not a constituent, and hence it is ‘invisible’ for purposes of licensing ellipsis. As a result, the only structure that is compatible with the partial VP-ellipsis shown in (11) is one in which the stranded adverbial (and its counterpart in the first conjunct) attaches above the elided VP. This allows the relevant string to be a constituent at the point when ellipsis is licensed, but it is incompatible with the left-to-right binding relation.

The contrast between VP-fronting and VP-ellipsis was perhaps the most notable empirical result of Phillips’ study, and it is the only one to have attracted counter-arguments. Importantly, one counter-argument that Phillips anticipated was the possibility that the contrast in (10-11) might simply indicate that movement and ellipsis are subject to different constraints. Phillips addressed this concern by showing that the same contrast observed between VP-ellipsis and VP-fronting is also found within a single construction – comparative ellipsis. Comparative ellipsis shows the same structural properties as VP-fronting when it shares the word order properties of VP-fronting, and it shows the same structural properties as VP-ellipsis when it shares the word order of VP-ellipsis. This is illustrated in (12) using the scope of durational adverbials. (12a) exhibits an ambiguity between a collective reading, which describes the total number of books read within a single one week period, and a distributive reading, which describes the number of books that were each individually read in less than a week. Phillips assumes that this scope ambiguity reflects a structural difference between high attachment of the adverbial (collective) and low attachment below the direct object (distributive). Both of these attachment sites are compatible with the word order in (12a), in which the adverbial follows the ellipsis site and hence should not interfere with the licensing of ellipsis. On the other hand, only the high attachment
needed for the collective reading is compatible with (12b), in which an adverbia is in both the antecedent and comparative clauses, and hence interferes with the constituency required to license ellipsis. This behavior of comparative ellipsis makes it unlikely that the contrast in (10-11) merely reflects quirks of ellipsis and movement. Unfortunately, this argument from comparative ellipsis has been overlooked in all of the criticisms of the left-to-right account of the contrast in (10-11).

(12) a. John read as many books as Bill did ___ in a week. [collective & distributive readings ok]
   b. John read as many books in a week as Bill did ___ in a month. [collective reading only]

The proposal that the results of different structural diagnostics reflect the dynamics of left-to-right structure building has much to recommend it. It addresses a fundamental problem in syntactic theory, it promises to bring together a diverse set of empirical facts under a general explanation, and it has interesting consequences for the mentalistic interpretation of grammatical theories. But it is important to ask whether the proposal has stood the test of time. The answer to this question appears to be a resounding “yes and no”.

As a general account of why structural diagnostics yield different results the account has fared well, in the respect that no new accounts have emerged that offer a similarly broad account of the problem. A few studies have offered accounts of specific phenomena in terms of the more mainstream bottom-to-top derivations, and we will have more to say about those in a moment, but all of these have had a restricted empirical focus. There has been an upsurge of interest in coordination and right-node raising (e.g., Wilder, 1999; Sabbagh, 2007; Grosz, 2009; Larson, 2012); there is much new research on ellipsis (e.g., Merchant, 2001, 2004; Elbourne, 2008; Johnson, 2009); there is a rich set of findings on the interaction of scope and binding with movement processes (e.g., Fox, 1999; Boeckx, 2001); and so forth. But to our knowledge there have been no general attempts to explain why structural diagnostics yield diverse results.

However, one could argue that the general accounts of structural diagnostics (by Pesetsky, Phillips, Steedman, and others) have fared poorly in the respect that the issue that they highlight has not been widely acknowledged as a challenge for syntactic theory. We are unaware of efforts to tackle the broader question that was offered as evidence for left-to-right derivations. Additionally, the account of structural diagnostics in terms of left-to-right derivations has made little progress in terms of new empirical coverage. In particular, little work has tested whether the proposal succeeds or fails when extended to languages with different structural properties. This probably reflects the fact that there is little broad interest in the typology of structural diagnostics.

The proposal that structure building proceeds in a left-to-right fashion has also fared well in the respect that many new proposals have emerged that offer accounts of various linguistic phenomena in terms of left-to-right derivations. These include a number of proposals framed within a similar transformational phrase-structure grammar approach to Phillips, covering such phenomena as locality of movement (Richards, 1999; Chesi, 2007; Zwart, 2009), subject-auxiliary inversion (Bruening, 2008), reconstruction asymmetries (Guillot, 2006), prosodic phrasing (Guimaraes, 1999; Wagner, 2005; Shiobara, 2008), expletive-associate relations (Richards, 1999), syntactic amalgams (Guimaraes, 2004), sprouting (Chung, Ladusaw, & McCloskey, 2011), right node raising (Park, 2005) and argument cluster coordination (Choi & Yoon, 2006), German ‘SLF’ coordination (Fortmann, 2005), and German adjective inflection (Schlenker, 1999). In addition, a number of interesting recent proposals have emerged in other grammatical frameworks in which left-to-right derivations play
a crucial role (e.g., Kempson, Gabbay, & Meyer-Viol, 2001; Cann, Kempson, & Marten, 2005; Shan & Barker, 2006; O’Grady, 2005).

On the other hand, proposals about left-to-right derivations have been relatively unsuccessful in the respect that there has been little genuine debate about the appropriateness of specific derivational orders, or indeed about the need for any derivations at all in syntax. Most views on these issues are currently relatively entrenched among syntacticians, and on all sides of the issue one encounters little interest in critically evaluating assumptions about how structures are put together. As a result the debate, such as it is, shows few signs of progressing towards a resolution.

We are aware of just a couple of proposals that aim to capture the generalizations that Phillips attributes to left-to-right derivations in terms of more standard bottom-to-top derivations. These have mostly focused on the account of the interaction between VP-fronting or VP-ellipsis with scope and binding (10-11). We will discuss these in a little more detail.

8. Scrambling in English (Baltin, 2003, 2006)

Baltin offers separate accounts of why partial VP-ellipsis bleeds binding possibilities (Baltin, 2003) and why partial VP-fronting does not (Baltin, 2006), within an approach that assumes bottom-to-top derivations and extensive scrambling of VP-internal phrases in English. Baltin argues that partial VP-ellipsis occurs when modifier phrases undergo string-vacuous rightward scrambling to a position above the VP, allowing the entire ‘remnant’ VP constituent to undergo deletion as a constituent. Given the additional assumption that binding constraints apply after scrambling, Baltin is able to capture the observation that material inside the elided VP cannot bind or take scope over elements inside the stranded modifier phrases (Baltin, 2003). This leaves Baltin with the challenge of explaining why fronting of partial VPs has a different effect than partial VP-ellipsis, such that elements inside the fronted VP are able to bind or take scope over stranded modifier phrases, as in (10). Previous analyses have taken it for granted that the binding relations must reflect a configuration in which the fronted VP is replaced (‘reconstructed’) in its underlying position, leading to Pesetsky’s paradox. In contrast, Baltin assumes that the binding/scope relations reflect the surface position of the fronted VP. Of course, if the VP is fronted as a constituent, then the sub-constituents of the VP should fail to c-command elements outside of the VP. Instead, Baltin assumes that the fronting process that appears to move a VP constituent is in fact a series of movement operations that independently front the verb and scramble the other sub-constituents of the fronted VP (Baltin, 2006). The fact that these operations yield a word order that looks exactly like a fronted VP constituent is entirely coincidental under this account. By assuming that each of the sub-constituents of the VP is fronted independently, Baltin is able to explain how they are able to c-command modifier phrases that are stranded by the fronting process, thereby capturing the lack of interference between movement and binding. Thus, whereas Phillips argues that the contrast between partial VP-fronting and partial VP-ellipsis in (10-11) reflects the linear order properties of the two processes, Baltin proposes that they differ because neither is a process that really targets a partial VP constituent. Ellipsis applies to a larger full-VP constituent, which only appears to be smaller because of string-vacuous scrambling. Meanwhile, fronting applies to smaller units separately, only appearing to target a single larger constituent because the different fronting operations conspire to recreate exactly the same word order that existed prior to movement. Everything is not the way it seems.

Baltin’s account of the interaction between VP-ellipsis and scope/binding (Baltin, 2003) has much in common with Phillips’ account, in the respect that both accounts claim that ellipsis can strand VP-modifiers because the stranded modifiers occupy positions higher than the elided constituent. In both cases this captures the observation that a stranded anaphor cannot have an antecedent inside the elided material. The main
difference between the two accounts is that Baltin assumes that the stranded modifiers move from an initial low position to a higher position, which is actually to the left of the original VP constituent, whereas Phillips proposes that the stranded modifiers are initially merged in a high position. Baltin offers no reasons why his account of the English facts should be preferred. He does contend that his account better captures a parallelism between binding possibilities in English VP-ellipsis and Dutch scrambling, but since the Dutch facts simply show that binding possibilities respect surface word order, we do not find this consideration to be decisive. However, Baltin’s article does present one simple generalization that does not follow from Phillips’ account. Baltin points out that although binding relations are not possible between the elided and stranded VP material, ellipsis does not interfere with binding relations between multiple stranded modifiers (13).

(13) Tom played his guitar for the boys at their graduation, and Bob did so for the girls on each other’s birthdays.

(13) illustrates the relatively mundane fact that the NP in the first stranded modifier *the girls* can bind a reciprocal in the second stranded modifier, just as it would if no ellipsis had occurred. This does not follow from Phillips’ account, since he assumes that stranded modifiers are right-adjointed to VP, with the consequence that modifiers on the right should be structurally higher than those on the left. This (embarrassing) discrepancy calls for a modification of Phillips’ account of the relative positioning of multiple stranded modifiers.

Baltin’s account of the lack of interference between VP-fronting and scope/binding relations (Baltin, 2006) contrasts more strongly with Phillips’ account of Pesetsky’s paradox. Phillips adopts the standard assumption that VP-fronting involves fronting of a constituent, and that scope/binding relations are based on the non-fronted (‘underlying’) position of the VP. He attributes the apparent conflict between VP-fronting possibilities and scope/binding facts to the dynamics of left-to-right structure building. In contrast, Baltin assumes that Pesetsky’s paradox arises in situations where a number of different movement operations proceed independently. First, the adverbials that appear to be stranded in sentences like (10) undergo scrambling out of the VP to a position where they ultimately appear to have not moved at all. Next, the remaining VP material is fronted as a constituent to a position below the subject NP (where none of that material ever surfaces). Following this, the arguments and modifiers that remain inside the VP are moved individually to positions above the subject NP, from where they c-command the rest of the clause to their right. Finally, the VP, which by this point contains only the verb, is moved to a Topic position at the front of the clause. This final step restores the linear order of the original VP constituent, although in Baltin’s analysis it is not a constituent that has been fronted.¹⁰

Opinions may differ on how compelling one finds this way of capturing Pesetsky’s paradox using standard bottom-to-top derivations. Importantly, though, Baltin offers empirical arguments that he claims favor his account over others. However, we think that these arguments do not ultimately favor his account.

¹⁰ Baltin appears to assume that there are other circumstances where superficially similar word orders are derived by genuine fronting of a VP constituent, leading to cases where material inside the fronted VP cannot c-command stranded material to its right. Hence, some sentences may involve simpler derivations, but crucially these simpler derivations are not intended to account for cases of Pesetsky’s paradox like (10).
First, Baltin offers a couple of arguments for analyzing the position of ‘stranded’ adverbials as being higher than the underlying position of the VP, contrasting with Pesetsky’s and Phillips’ assumption that the binding possibilities in (10) must reflect a right branching structure in which the stranded PP attaches low within the VP, such that it is c-commanded by the direct object NP. Baltin points to restrictions on British English do anaphora, a phenomenon that has much in common with the do so VP-anaphora found across many varieties of English. British English do anaphora, just like its do so counterpart, can strand true non-argument adverbials (14a), but cannot strand argument PPs (14b).

(14) a. Although he wouldn’t visit Sallyi on her birthday, he would do ___ on heri anniversary.

    b. *Although I won’t put the book on the table, I will do ___ on the mantelpiece.

Based on this and other properties of the construction, Baltin argues that the adverbials must be attached in a position higher than the rest of the VP. This is a reasonable analysis of the position of adverbials in VP-ellipsis, and in fact it is the same as the analysis of partial VP-ellipsis that Phillips (1996, 2003) adopts. However, since the challenge is to explain why the scope/binding properties of stranded adverbials differ between VP-fronting and VP-ellipsis, Baltin’s argument from do anaphora, a form of VP-ellipsis, cannot be assumed to generalize to VP-fronting. In Phillips’ account, it is the position of the stranded adverbial that differs between VP-fronting and VP-ellipsis, and in Baltin’s account it is the position from which object NPs bind into the stranded adverbials that differs. The argument from do anaphora cannot decide between these two alternatives.

Baltin offers a further argument for the high position of adverbials stranded by VP-fronting, and this argument more clearly challenges the left-to-right account of Pesetsky’s paradox. Baltin observes that examples like (15a), in which a quantificational direct object inside a fronted VP appears to license bound variable pronouns inside the stranded adverbial phrase, are exactly the kinds of phenomena that motivated Pesetsky’s paradox. He points out that the variant in (15b), where the adverbial contains additional VP-ellipsis, creates a possible instance of antecedent contained deletion (ACD) if the adverbial is attached low inside the VP, as is assumed in Phillips’ account. Phillips argued that low attachment of an adverbial inside a VP destroys the constituency of the verb and the direct object, making them invisible to subsequent syntactic processes. By this logic, the VP-ellipsis inside the adverbial in (15b) should not be able to take the VP visit every prisoner as its antecedent. Baltin argues that if the ellipsis is, in fact, possible, and if it does not interfere with establishing a bound variable interpretation for the pronoun, then this implies that the binding relations must reflect a higher position for the adverbial, and also a correspondingly higher position for the antecedent of the bound variable pronoun.

(15) a. Visit every prisoner, though I may ___ after his, lawyer visits him, , it won’t matter.

    b. Visit every prisoner, though I may ___ after his, lawyer does ___, it won’t matter.

Cases like (15b) certainly aim at the heart of what Phillips’ account seeks to capture, by challenging the generalizations about the interactions between fronting, ellipsis, and scope/binding. However, we suspect that more work is needed to clarify the empirical facts in this domain before firm conclusions can be drawn. The status of (15b) and sentences like it strike us as unclear, and in related constructions the facts do not clearly favor Baltin’s conclusion. The examples in (16) contain at most one bound variable pronoun, and offer the potential for a scope ambiguity between a collective reading, where the subject finished reading the set of books before Sally
did, and a distributive reading, where the subject read each individual book before Sally read that book. To the extent that (16b) is possible at all, the collective reading seems to be more readily available than the distributive reading, in contrast with (16a), where the singular pronoun forces the distributive reading. This suggests that VP-ellipsis may indeed change the structural position of the adverbial phrase, contrary to Baltin’s argument. However, we should emphasize that Baltin’s observation highlights an interesting domain where more work is needed.

(16) a. Read every book, though he may before Sally reads it, she’ll still get the better grades.

       b. Read every book though he may before Sally does ___, she’ll still get the better grades.

Baltin also argues that his account is supported by striking correlations between constraints on pseudogapping and constraints on Pesetsky paradox effects. According to his account, both pseudogapping and the partial VP-fronting that create Pesetsky paradox effects reflect scrambling operations that remove arguments from a VP. Thus, any non-trivial correlations between these two processes would provide striking support for his parallel analysis. Baltin’s arguments are based on examples of pseudogapping and VP-fronting that involve complex VPs with more than one verb. (17) shows that complex predicates involving an infinitival under a subject control predicate allow both processes.

(17) a. Although I didn’t try to persuade Sally, I did ___ Martha.

       b. Try to visit every prisoner, though I may ___ after his,i lawyer does ___, I’m not sure they’ll be successful.

In examples with more complex predicates pseudogapping is clearly degraded. (18a) contains a main clause object control verb, such that an overt NP appears between the two verbs, and the pseudogapped clause does not allow a construal in which the gap corresponds to persuade Sally to visit. Baltin argues that although the VP-fronting in (18b) is possible, it does not allow the indicated bound variable relation from the fronted direct object to the stranded adverbial, i.e., Pesetsky’s paradox effects do not obtain. Similarly, Baltin points out that complex pseudogapping is impossible with predicates that select a non-finite complement headed by from (19a), and claims that when the same predicates undergo VP-fronting they also fail to show Pesetsky paradox effects (19b). Baltin argues that both of these restrictions are due to constraints on the scrambling processes that are part of his analysis of both constructions. Moreover, Baltin argues that the examples in (18c, 19c) that use coreferential pronouns instead of bound variable pronouns are acceptable, since coreference does not require c-command. He claims that these contrasts are mysterious under either Pesetsky’s or Phillips’ accounts.

(18) a. *Although I didn’t persuade Sally to visit Martha, I did ___ Susan.

       b. *Persuade Sally to visit every student, though I may ___ on his,i graduation day, it won’t matter.

       c. Persuade Sally to visit Tom, though I may ___ on his,i graduation day, it won’t matter.

(19) a. *Although he didn’t refrain from visiting Martha, he did ___ Susan.

       b.*Refrain from visiting every student, though she may ___ on his,i graduation day, it won’t matter.
c. Refrain from visiting Tom, though she may ___ on his graduation day, it won’t matter.

These correlations are impressive if they are accurate, but we suspect that the restrictions are broader than Baltin suggests, and do not reflect constraints on scrambling. Baltin is probably correct that pseudogapping cannot target sequences that include an overt subject NP, as shown by the more minimal contrast in (20a-b). However, the degraded status of (20b) is probably not specifically due to the subject NP that intervenes between the main verb and the infinitival, since examples with intervening adverbials are similarly degraded (20c).

(20)  
a. Although I didn’t want to visit Martha, I did ___ Susan.

b. *Although I didn’t want Sally to visit Martha, I did ___ Susan.

c. *Although I didn’t try as hard as I could to visit Martha, I did ___ Susan.

The contrasts in binding possibilities between (17b) on the one hand and (18b) and (19b) on the other hand are, frankly, too subtle for us to reliably assess. However, we suspect that the binding and coreference facts that Baltin points to are reflections of a more general constraint that affects VP-fronting and VP-ellipsis alike, and that was already discussed in Phillips (1998), an earlier version of Phillips (2003). In any sequence consisting of Aux [VP null] adverbial the adverbial must be interpreted as a clausemate of the auxiliary. This can be seen clearly in the examples in (21), where the locative PP in the basement must be understood as modifying the matrix verb resolve rather than the embedded verb fix. The verb resolve is preferable to other control verbs like want and try used in some of the examples above, since the time and location of want and try and their complements coincide, a problem that does not arise with resolve and its complement.

(21)  
a. Wallace resolved to fix the motorcycle in the garage, and Gromit did in the basement.

b. Wallace needed to resolve to fix the motorcycle, and resolve to fix the motorcycle he did in the basement.

Phillips (1998) characterizes the clausemate constraint on modifier construal illustrated in (21) as an independent constraint that is unrelated to the dynamics of left-to-right structure building. Here we offer no deeper explanation of the constraint, although we note that it is not an unnatural constraint. We suggest that it is this constraint that is responsible for the contrasts that Baltin presents as a challenge for the left-to-right account.

Baltin’s attempt to reconcile traditional bottom-to-top derivations with constituency conflicts faces a number of further challenges. First, Baltin’s account of the contrast between VP-ellipsis and VP-fronting is based on construction-specific differences between the two phenomena, an approach that is challenged by the similar contrasts found internal to comparative ellipsis. Second, if Baltin is correct that VP-fronting is

---

11 The 1998 manuscript includes some sections that were removed in the published version to satisfy length constraints.
12 Landau (2007) shows that the clausemate constraint applies even with predicates like want and try, by using frequency adverbs, which disambiguate the main clause and embedded clause interpretations, as in (i).
(i) John intended to try to meet Mary often, but try to meet Mary he did only rarely. [rarely = try]
not a unitary operation, and that fronted arguments of the verb bind stranded modifiers from their fronted position, then we should expect that the same phrases should be able to bind subject NPs, a clearly inaccurate prediction (22). Third, Baltin’s account of binding in VP-fronting is at odds with the well-supported independent generalization that fronted predicates obligatorily reconstruct for purposes of binding (Huang, 1993; Heycock, 1995; Takano, 1995). Fourth, it is puzzling under Baltin’s account (as he acknowledges) that pseudogapping should allow argument stranding whereas VP-fronting does not (23). More generally, as argued by Landau (2007) partial VP-fronting is attested in many languages that lack productive scrambling rules (e.g., Italian, Spanish, Brazilian Portuguese, Hebrew), and partial VP-fronting in these languages allows stranding of phrases that cannot otherwise be removed from the VP. Finally, it is hard to overlook the fact that fronted VPs really do look like fronted constituents: they coordinate, they tolerate no interruptions, etc. This suggests that they really are constituents.

(22) a. *Visit every prisoner; though his lawyer may __ before the hearing, it won’t matter.
 Centro de Linguística da UFMG

b. *Praise the girls; though each other’s mother did __, they still felt sad.

(23) a. Bill visited Sally, and Bob ___ Martha.

b.*Visit though he may ___ Sally, she won’t be satisfied.


Lechner (2003) offers a bottom-to-top derivational account of the contrasting properties of VP-fronting and VP-ellipsis that, like Baltin’s analysis, assumes that partial VP-fronting in English relies on counterparts of the scrambling and remnant VP-movement operations found in German. Lechner proposes that there are construction-specific differences in the way that modifiers are combined with the predicate in VP-fronting and VP-ellipsis constructions. In VP-fronting a modifier can be initially attached low in the VP, then later be removed from the VP by scrambling, clearing the way for partial VP-fronting, and can finally be reconstructed to its original site for purposes of evaluating binding relations. In VP-ellipsis, on the other hand, Lechner claims that the strategy of scrambling and then reconstructing the stranded modifier is not available, due to economy conditions that force the stranded modifier to directly merge in a position higher than the rest of the predicate. This high merger site accounts for why it is not possible for elided VP material to bind into the stranded modifier.

Since Lechner takes the position that there are construction-specific differences between partial VP-fronting and partial VP-ellipsis it is not clear how his account could capture the contrast within comparative ellipsis described above. A more detailed critique of Lechner’s approach can be found in Landau (2007). More relevant to our current concerns, however, is to assess the empirical challenges that Lechner presents for the left-to-right structure building account of VP-fronting and VP-ellipsis.

Lechner points out that although partial VP-ellipsis blocks binding from an elided object into a stranded adverbial (9), no such restrictions are observed in the corresponding phrases in the first conjunct of a VP-ellipsis construction (24). This is unexpected under the assumption that VP-ellipsis must be licensed by a constituent in the antecedent clause. Similar examples were already discussed in Phillips (1996), where it was acknowledged that the judgments do not directly follow from the left-to-right structure building account.

(24) John gave the books to them, on each other’s birthdays, and Mary did ___ on their first day of school.
Lechner also points to scope freezing effects in partial VP-fronting that do not follow from the left-to-right account. In (25) the stranded PP obligatorily takes wide scope relative to the fronted predicate, although Phillips’ account predicts that it should be possible to merge it in a low position inside the predicate, once the predicate is copied to its thematic position.

(25) a. David planned to deliver every book to one of the students, $\forall > \exists, \exists > \forall$

b. …[VP and deliver every book] he did ___ to one of the students.*$\forall > \exists, \exists > \forall$

More can be said about the details of these observations (see Phillips, 2002; Lechner, 2002 for discussion), but we suspect that the ultimate source of these effects lies in incremental interpretation. Once the binding relation in the first clause of (24) is established, it cannot be retracted based on the constituency requirements of the VP-ellipsis construction. The fronted predicate in (25) is interpreted in its fronted position, such that subsequent phrases must take scope outside this interpretive unit. These suggestions are certainly speculative, and they beg a number of questions, but we suspect that it is considerations such as these, rather than raw constituency constraints, that are ultimately responsible for the unexpected interpretations.

Lechner’s third criticism of the left-to-right approach parallels one of Baltin’s concerns. He points to examples of antecedent contained deletion (ACD) in a phrase stranded by VP-fronting (26), and argues that these should be impossible under the left-to-right structure building analysis. The problem is that if the stranded phrase merges low inside the fronted predicate, then the antecedent for the ellipsis site contains the ellipsis site, leading to an infinite regress. It is important that these cases involve stranded (but optional) argument phrases, as these are not assumed to have the option of simply merging high in the VP to avoid the constituency problem.

(26) Mary asked him to deliver a book to some of the boys, and [VP deliver a book] he did ___ to every boy Mary wanted him to ___.

The second ellipsis site in (26) is interpreted roughly as deliver a book to $x$, where the variable is bound by the quantificational NP every boy. Lechner is correct that the availability of this antecedent for ellipsis does not follow directly from the left-to-right structure building account. In fact, the puzzle that such cases present can be seen in simpler examples that lack VP-fronting, such as (27).

(27) Harry delivered a book to every boy Mary wanted him to ___.

We should emphasize that the puzzle that such examples present appears to be specific to ACD. In other instances of VP-ellipsis, judgments are as expected. The examples in (28) show VP-ellipsis inside a relative clause, as in ACD, but the head of the relative clause is non-quantificational, and the ellipsis site is marked by do so. The ellipsis is fine in (28a), where the ellipsis antecedent read a book is a constituent that does not include the ellipsis site. The ellipsis is impossible in (28b), where the intended ellipsis antecedent contains the ellipsis site. Since regular VP-ellipsis behaves as expected, and ACD in examples like (26-27) creates exactly the same constituency puzzle for the left-to-right approach that it creates for traditional approaches, we assume that cases like (26-27) should be handled in the left-to-right approach using similar mechanisms to those used to handle ACD in other approaches (e.g., quantifier raising).

(28) a. John read a book at a table where Mary had done so ___ half an hour earlier.

b. *John put a book on a table where Mary had done so ___ half an hour earlier.
Lechner’s final concern involves the clausemate constraint on the construal of phrases stranded by VP-ellipsis and VP-fronting. He notes that the constraint amounts to a restatement of the facts, a criticism that is largely accurate. There is nothing inherently wrong with the fact that the generalization is not a direct consequence of left-to-right structure building, since we should not expect derivational order to be a cure-all that explains all generalizations. But it is, of course, interesting if the generalization turns out to follow from independent principles. Lechner offers one such proposal, which essentially treats the clausemate constraint as a consequence of locality restrictions on extraposition processes. Landau (2007) provides empirical arguments against Lechner’s interpretation of the restriction, and proposes an alternative account of his own. In our impression, the jury remains out on the appropriate analysis of the clausemate constraint.

10. Late merger of adjuncts (Landau 2007)

Landau (2007) offers an interesting discussion of constraints on partial VP-fronting, with a focus on the requirement that fronted VPs be at least partially complete VP constituents (Phillips, 2003). Landau discusses data from English and from Hebrew, a language that appears to be particularly well suited to examining partial VP-fronting. His account of Pesetsky’s paradox is framed in terms of bottom-to-top derivations, but his account has much in common with the left-to-right structure building account. Landau resolves the paradox by assuming that VP-fronting occurs prior to merger of stranded adjunct phrases into the null VP-copy. Thus the fronted VP is a constituent at the point where movement occurs, although it is no longer a constituent in its underlying position once the stranded adjunct is inserted into the structure. This order-based solution is very similar to Phillips’ account, although the two proposals differ in a number of details.

Landau criticizes the left-to-right structure building account on the grounds that it does not offer a compelling account of the requirement that fronted VPs be potentially complete VPs. Phillips (2003) did offer a half-hearted explanation, and Landau does a good job of demonstrating the limitations of that account. Landau offers his own account of the restriction, based on a requirement that verbs’ thematic requirements be satisfied locally. If Landau is correct, then his account of the constraint could straightforwardly be implemented in a wide variety of grammar formalisms, independent of derivational order, and so it could be readily incorporated into the left-to-right account.

Landau offers argues that his approach is supported by an interesting additional generalization, based on an extension of well-known asymmetries in reconstruction effects pointed out by Freidin (1986) and Lebeaux (1988). Freidin and Lebeaux pointed out that clausal arguments of fronted NPs show reconstruction effects for purposes of Condition C (29a, 30a), but relative clauses do not. If the entire fronted NP in (29b, 30b) is interpreted in direct object position, then the indicated coreference relation should be ruled out by Condition C, yet coreference is possible.

(29)  
   a. ?*Which argument that Johni is a genius did hei believe?
   b. Which argument that Johni made did hei believe?(Lebeaux, 1988)

(30)  
   a.?*The overwhelming evidence that Henryi was a spy, hei refused to accept.
   b. The overwhelming evidence that Henryi had amassed, hei refused to present.
A popular approach to this paradigm is to assume that relative clauses and other adjuncts are able to escape the effects of Condition C because they are merged with the wh-phrase in its fronted position, and hence are never c-commanded by the coreferential pronoun. Landau offers an interesting twist on this generalization, arguing that there is a correlation between the phrases that can escape Condition C effects when they are fronted and the phrases that can be stranded when a predicate is fronted. This correlation is expected under Landau’s approach, since both effects are attributed to late merger of the non-obligatory phrases. Landau extends his account to cover an additional contrast between different types of modifiers inside fronted predicates. Although a relative clause inside a fronted nominal is able to escape Condition C effects (31a), the Condition C effect reemerges when the same nominal is embedded inside a fronted predicate (31b; Takano, 1995). In contrast, Landau observes that Condition C can be avoided when the critical noun is part of a VP-modifier in a fronted predicate (31c). Landau argues that this can be explained by his account, if late adjunction of adjuncts is further constrained such that it cannot apply inside predicates. The modifier containing the noun Mary is inside the fronted predicate in (31b) but is presumably adjoined to the top of the fronted predicate in (31c). Landau suggests that the constraint on late adjunction inside predicates could also explain the clausemate condition on the interpretation of stranded modifiers.

(31)  
a. Food that Maryi cooks, shei knows I would never eat.  
b.*Eat food that Maryi cooks, shei knows I never would.  
c. Eat food at Maryi’s party, shei knows I never would.

Interestingly, the contrasting possibilities for escaping Condition C effects may be equally compatible with a left-to-right structure building account, without a need to invoke late adjunction processes. Phillips (1998; section removed in 2003 version) offers an account of the Freidin-Lebeaux facts in (29-30), proposing that the contrast simply reflects constituent structure. If it is assumed that a relative clause is adjoined to NP, then there is a smaller constituent (e.g., which argument) that may undergo reconstruction, without reconstructing the critical name and inducing a Condition C violation. On the other hand, if clausal complements of nouns are sisters of N, then there is no constituent that includes the wh-phrase which argument and excludes the complement clause, and consequently the entire nominal must reconstruct, leading to a Condition C violation. This account straightforwardly extends to the contrast in (31ab). (31a) is structurally identical to (29b, 30b), and hence the relative clause may avoid reconstruction. But once the nominal is embedded inside a predicate (31b) there is no constituent that includes the verb and the direct object and excludes the relative clause, and thus the entire predicate must reconstruct.

Several authors have questioned the accuracy of the Freidin-Lebeaux contrast (e.g., Bianchi, 1995; Kuno, 1997; Postal, 1997; Lasnik, 2003). These authors have pointed to examples like (i-iii), which suggest that extraction of arguments of NP may sometimes also bleed Condition C.

(i)Which piece of evidence that Johni was guilty did hei successfully refute? (Lasnik, 2003)  
(ii)The claim that the directori was corrupt, hei was unwilling to discuss. (Postal, 1997)  
(iii)Whose claim that the senator i had violated the campaign finance regulations did he i dismiss as politically motivated? (Kuno, 1997)

The facts in (i-iii) are interesting, but we do not think that they undermine the interest of the Freidin-Lebeaux contrast. Adjuncts to NP appear to reliably escape Condition C effects, whereas there is some variability regarding the status of arguments of NP. Landau (2007) suggests that this may be because some finite complements that have been analyzed as nominal complements should, in fact, be treated as adjuncts (following Stowell, 1981; Safir, 1999).
The status of reconstruction for VP-modifiers is less clear. Landau argues based on examples like (31c) that such modifiers can freely escape Condition C effects. But others, going back as far as Hasegawa (1983; observation attributed to Joan Bresnan) have argued that fronting of predicates can never bleed Condition C. Phillips (1998) offers the examples in (32) of VP-modifiers, which appear to be subject to obligatory reconstruction.

(32)  a. *[Playing cards until long after Lucy’s bedtime] though she was, she, was not at all tired the next morning.

b. *[Read a biography in Wallace’s living room] though he would like to, there’s no chance that he, actually will.

The modifiers that undergo obligatory reconstruction in (32) are modifiers that would normally allow stranding in cases of VP-fronting (33). Such cases cast doubt upon Landau’s correlation between modifiers that escape Condition C and that can be stranded by predicate fronting. More generally, though, the contrast between (31c) and (32) suggests that more work is needed to understand under what circumstances VP-modifiers can escape Condition C effects.

(33)  a. Play cards though Lucy did ___ until long after her bedtime, she was not at all tired the next morning.

b. Read a biography though Wallace did ___ in his living room, he still couldn’t relax.

Summarizing, a number of specific challenges have been raised for the account of constituency conflicts offered by Phillips (1996, 2003), framed in terms of more ‘conservative’ approaches to syntactic derivations, i.e., derivations that start at the end of a sentence and end at the beginning. Some of these challenges can be addressed easily, while others are less straightforward. In some cases, such as Landau’s account of the potential complete VP constraint, interesting new generalizations could be straightforwardly grafted onto the left-to-right account. We are unaware of any alternative proposals that have the scope of the left-to-right account, or that offer such a compelling bottom-to-top analysis to motivate abandoning the left-to-right account.

11. Conclusion

Discussions of left-to-right order in syntactic derivations tend to focus on standard considerations of descriptive adequacy – whether left-to-right order provides broader coverage of (un)acceptable sentences. But the elephant in the room in such discussions is the possibility of understanding human grammatical computation in more realistic terms, as mental processes that operate in real time and that are open to scrutiny using a host of cognitive and neuroscientific tools. Here we have argued that current evidence favors the view that human grammatical abilities are best understood as a single structure building system that assembles syntactic structures in a (roughly) left-to-right order, and is a key component of systems for parsing and production. This position departs from widespread assumptions in linguistics in two ways. For at least 40 years it has been standard to assume that human grammatical competence is an implementation independent system, and that this system assembles sentences via bottom-to-top derivations that typically proceed in the opposite order than normal comprehension and production. We have argued that there is very little evidence for the implementation independence of human grammatical abilities, and that the motivations for bottom-to-top syntactic derivations are less than overwhelming. Meanwhile, the feasibility of the real-time approach to grammatical computation has been reinforced by linguistic and psycholinguistic evidence of recent years. This implies a research program for grammatical theory that goes far beyond the traditional concern with classifying acceptable and unacceptable sentences, and it suggests that
grammatical theories should be accountable to a much richer body of evidence, particularly evidence on the time course of grammatical processes. This is just what we should want.

References


Treegrowth Dynamics*

RUTH KEMPSON\textsuperscript{a}, RONNIE CANN\textsuperscript{b}, and LUTZ MARTEN\textsuperscript{c}

\textsuperscript{a} Department of Philosophy, King’s College London
\textsuperscript{b} School of Philosophy, Psychology and Language Sciences, University of Edinburgh
\textsuperscript{c} Departments of Linguistics and Languages and Cultures of Africa, SOAS, University of London

\texttt{a} ruth.kempson@kcl.ac.uk
\texttt{b} ronnie@ling.ed.ac.uk
\texttt{c} lm5@soas.ac.uk

In this paper we address the challenge of clitic clusters in Romance languages and the puzzling gaps in the clitic templates which they display, and show how the distribution of clitics and these apparent gaps are grounded in one and the same structural restriction: a restriction on tree-growth. This parsing-inspired restriction states that in building up structurally underspecified relations, only one such relative weak relation can be constructed at a time. This is a core restriction underpinning concepts of tree growth central to the Dynamic Syntax framework which argues that natural-language syntax is grounded in tree growth mechanisms that reflect on-line processing dynamics. We apply this constraint to explain the restriction on the morphological clitic templates of Romance currently referred to as the Person Case Constraint (PCC). We show that the explanation of such gaps in terms of preclusion of more than one structurally weak relation is grounded in the diachronic calcification of tree-growth strategies that had earlier constituted a freely available set of options for building up interpretation via flexible word orders, individual clitics and clitic clusters severally displaying the various strategies. Apparent counterexamples to the constraint are explained in terms of the availability of an alternative adjunct strategy not involving any such underspecification which, in those languages apparently violating the constraints, had led to homonymy in the clitic system. Finally, the force of the structural basis for the PCC is buttressed by the demonstration of its applicability to explain the object-marking puzzle of Otjiherero, a Bantu language which, despite allowing construal of object agreement markers as both indirect and direct object, never allows both to be co-present.

* This paper was first presented in 2008 and there are a number of people who have contributed to these ideas and to subsequent developments that have arisen from it. First there is Miriam Bouzouita, without whom the Dynamic Syntax account of syntactic change might never have got off the ground. We are very grateful to her for letting us use her data, for her insights into methodologies for diachronic study and the nature of diachronic developments of Spanish clitics. There has also been substantial collaborative work done on clitic cluster distributions by the first author and Chatzikyriakidis (see Chatzikyriakidis 2010, 2011, Chatzikyriakidis and Kempson 2011). These two colleagues in particular deserve a special vote of thanks. We are also grateful to Jekura Kavari for supplying the Otjiherero data and for discussion of their analysis with the third-named author. There is also Eleni Gregoromichelaki, Jieun Kiaer, Andrew Gargett, and Wilfried Meyer-Viol to thank for input as these ideas have emerged. This work was reported at a Person Case Constraint workshop in November 2007 at Queen Mary University of London, and we thank participants for comments. Normal disclaimers apply.

© Copyright 2013 Ruth Kempson, Ronnie Cann, Lutz Marten
STIL – Studies in Linguistics. 6:49-81
Special Issue on “Directionality of Phrase Structure Building” ISSN: 2281-3128
1. Preliminaries
In this paper, we address the challenge of explaining clitic clusters in the Romance languages, with the puzzle of apparent Person Case Constraints that are associated with these clusters which, though very widespread, by no means universally hold, with variation between otherwise very closely related languages. Such clitic clusters are often characterised as irreducibly opaque morphological templates, little more than a cleaned-up disjunctive statement of the facts of the matter. We argue that these clusters and the gaps in the paradigms associated with them can be explained in terms of a constraint on procedures for building up representations of content in real time, a constraint which also serves to determine limits on NP-placement in free scrambling languages. Our case study is the clustering properties displayed in the Romance languages, our principal focus being on Latin and Medieval Spanish. What we show is that the various types of clitic behaviour match the range of strategies available for the structure-building mechanisms underpinning construal of sequences of NPs in Latin, with its free scrambling of NPs before the verb; and we argue that the emergence of the clitic systems of Romance was a calcification of these strategies variously associated with individual lexical specifications. We go on from this to demonstrate that the gaps in these paradigms, in particular the so-called Person Case Constraint, can also be explained in terms of the very same constraint that underpins scrambling of NPs. So the explanation both of clitic clusters and their constraints is essentially syntactic, but with a twist, as the structural explanation to be provided makes essential use of the growth of representation of content to be attributed to the string, so essentially semantic.

What we argue is that mechanisms for inducing structural representations of content involve constructing structurally underspecified relations, in particular for structural relations before a verb is processed; and, as we shall show, all natural language systems are defined so as to preclude the building of more than one such weak structural relation at a time. This, we claim, is a core structural restriction underpinning all tree growth mechanisms in natural language, a property of the tree logic in which such mechanisms are grounded. In the case of the clitics, the restriction is largely but not exclusively realised by the structurally syncretic case forms, and we argue that the syncretic morphology accurately depicts the structural lack of specification in the tree-relation constructed. Apparent counterexamples can be explained in terms of the availability of an alternative parsing mechanism that does not involve such a structurally underspecified relation which, through its widespread use in those languages apparently violating the constraint, became encoded as a homonymous clitic form. Independent evidence of the form of explanation is demonstrated by the applicability of the same procedural tree-growth restriction to explain a gap in the paradigm of object marking in Otjiherero, a Bantu language allowing object marking across a range of different complement types such as direct, indirect and locative objects, but only ever allowing one object marker in ditransitive constructions, irrespective of which grammatical relation is expressed through object marking. We will conclude that the broad applicability of the constraint of only one type of underspecified relation at a time within both syntax and morphosyntax is strong evidence of the parsing grounding of structural properties of language, hence that core mechanisms of grammar should reflect processing dynamics.
1.1 The Person Case Constraint: A sketch

The Person Case Constraint is a label for a puzzling pair of restrictions in morphosyntax. In languages that have pronominal clitic systems, these commonly cluster together either occurring immediately preverbally, or in some not very clearly definable second position. Over and above the strictly syntactic problem of how to characterise this position, these clusters display such heterogeneity in the internal orderings which individual languages and/or dialects license that their variability is said to warrant a separate morphology component within the grammar. Nonetheless, they are subject to restrictions that hold with striking regularity across different language families. Though some gaps are random, others never, or almost never, occur. The most well-known of these has come to be known as the Person Case Constraint, of which there are two variant forms. The most widespread is that construal of either first and second person forms as an accusative can never co-occur with third person dative, a restriction which holds in all the Romance languages as long as we keep distinct the various usages of the dative clitic as ethical, reflexive, and impersonal. Thus, though there are occurrences of first or second person construed as dative followed by third person accusative in Medieval Spanish, just as in other Romance languages, as in (1), there is no reported occurrence of construal of either as accusative, with dative third person (data from Granberg 1988: 176):

(1) agora quiero **uos** lo descubrir             [Medieval Spanish]
    now want.1sg you.dat it.acc reveal.inf
    ‘Now I want to reveal it to you.’

The lack of first or second person accusative forms with a third person dative is a puzzling gap if such clitic forms are thought to be a mere listing of possible morphological forms. There is also a further restriction displayed by a subset of these languages: first and second-person pronominal forms should not co-occur, as in the French example in (2). In this case, the restriction is less definitive, and in Latin (3) and Spanish (4) such examples are wellformed, a problem which we return to:

(2) *Il me t’ a fait montrer un livre              [French]
    he 1sg 2sg have.3sg made show.inf a book
    ‘He made me show you a book.’

(3) qui **me tibi** fecerit hostem             [Latin]
    who.sg.nom 1sg.acc 2sg.dat make.3sg.perf enemy.sg.acc
    ‘who would make me an enemy to you.’ (Lucan *De Bello Civile* 1)

(4) **Te me** recomendaron                     [Spanish]
    2sg.acc 1sg.dat recommend.3pl
    ‘They recommended you to me.’

Such gaps, and indeed the less common lack of any such gap in some languages, are mysterious, since they appear to lack syntactic, semantic, or phonological explanation (Anderson 2005, Monachesi 2005). That this is neither a semantic or pragmatic restriction can be shown by the fact that in all these languages the restriction can

---

1 We list examples with clitic pronouns in bold for clarity.
simply be side-stepped by realizing one of these arguments as a strong pronoun or full NP (data from Rivero 1997):

(5) Si de nos te non partes … [Medieval Spanish]
    if from us you.refl not depart
    ‘If you do not leave us …’

In the face of the various challenges which these clustering facts present, many authors settle for a discrete morphological form of explanation, not characterising the phenomenon within the syntactic domain at all (see e.g. Nevins 2007, Rivero 2007). In this paper, however, we explore a perspective in which these restrictions can be seen as a consequence of a strong structural restriction, one that is a constraint on growth of interpretation. The restriction is a constraint debarring more than one underspecified structural relation at any one point in the unfolding construction process of establishing interpretation. We shall argue that the reason it emerges as a morphosyntactic restriction in the Romance clitic systems is that the patterns which Romance clitics severally display each constitutes a calcification of some processing strategy that was freely operative in the earlier Latin system, subject to general constraints on growth of interpretation, so the debarred clitic cluster sequences currently known collectively as the Person Case Constraint (PCC) (e.g. Adger and Harbour 2007) cannot arise because the earlier pattern of NP sequencing was precluded. The difficulty for any such structural account of this constraint is that some languages apparently fail to display PCC effects. So making this move will impose the challenge of explaining the apparent exceptions, but that too we will argue can be seen as a consequence of adopting a parsing perspective. The account as given here, though within a formal framework, is largely informal (see Bouzouita 2008a, 2008b, 2011, Bouzouita and Chatzikyriakidis 2009, Cann and Kempson 2008 for detailed lexical specifications, and Chatzikyriakidis and Kempson 2011 for a detailed account of PCC effects in an array of Greek dialects).

2. Background: Latin scrambling puzzles and the emergence of pronominal clitic systems
The background against which this account is set is the challenge posed by clitics, whose positioning and clustering behaviour are problematic, given current theoretical assumptions.

2.1 Clitic mysteries
Pronominal clitics are typically weakened quasi-affixal, quasi-pronominal devices, with a characteristic preference for occurring at some relatively early position in a finite clausal sequence, in some languages immediately following some first constituent or word (as seen in the Baltic languages), in other languages immediately preceding the finite verb (most Romance languages, including Modern Spanish); and yet others with some mixture of the two (Medieval Spanish, Cypriot Greek). The first of these alternatives, the second-position clitic placement, is hosted by a heterogeneous set of categories, commonly including complementisers, wh-expressions, negation markers, focused expressions, relative pronouns, verbs (if
nothing else precedes), and in some cases conjunction markers (see Bouzouita 2008a, 2008b, 2011 for more details for Medieval Spanish):2

(6)  ... **quien te algo prometiere ...** [Medieval Spanish]  
    who you.dat something would.promise  
    ‘the one who would promise something to you.’

(7)  **Quant le connocio Abdias homillo-s-le**  
    when him.acc recognised.3sg Abdias lowered-himself.refl-him.dat  
    ‘When Abdias recognised him, he bowed for him.’

(8)  **Que te dixo Heliseus?**  
    what you.dat said.3sg Heliseus  
    ‘What did Heliseus tell you?’

(9)  **Non los destroyré [...]**  
    not them.acc will-destroy.1sg  
    ‘I will not destroy them [...]’

(10)  **.ij. mios fijos te dexaré [...]**  
    two my sons you.dat will.leave.1sg  
    ‘My two sons, I will leave you [...]’

(11)  **Con aquellas se aiunto Salomon [...]**  
    with those himself.refl joined.3sg Salomon  
    ‘With those women, Salomon slept [...]’

(12)  **Oyo-l Ruben [...]**  
    heard-3sg.him.acc Ruben  
    ‘Ruben heard it [...]’

This set of environments resists any unitary syntactic characterisation, upon conventional assumptions. The puzzle, from a syntactic perspective, is what this array of variation can be grounded in?

Construal of clitics is also puzzling, as, for some clitics, their argument-role is fully determined by their form, but for others it is not. Accusative clitic pronouns in Romance for example are standardly relatively clear-cut in indicating a direct-object argument. Dative pronouns on the other hand invariably display a flexibility in construal which is displayed also in Latin, where the Greco-Latin tradition describes dative NPs as dividing into at least ten distinct types, from the marking of direct and indirect objects through to widely varying semantic uses, including possession, advantage, result and ‘interest’ (reported in van Hoecke 1996). Latin examples include:

2 In the main, illustrations are from 13th century Medieval Spanish and are taken from a corpus of Medieval Spanish collected by Miriam Bouzouita culled from the Fazienda de Ultramar, which dates from around 1230. All Medieval Spanish examples given are from this text unless stated otherwise.
(13) an *tibi* quisquam in curiam venienti
Q you.dat anyone.nom.sg into senate-house.acc.sg coming.dat.sg
assurexit? [Latin]
get-up.3sg.perf
‘Did anyone get up for you (to your benefit) when you came into the senate
house?’ (Cicero *In Pisonem*. 26)

(14) *quid mihi* Celsus agit?
what me.dat Celsus.nom.sg do.3sg.pres
‘How, pray, is Celsus?’ (Lit. ‘What to me Celsus does?’; Horace *Epistolae* 1, 3, 152)3

Following this Latin usage, the first and second dative clitic pronouns in the Romance
languages are commonly associated with a large number of distinct construals, the
particular range varying from language to language, and even from environment to
environment. For example, first and second person clitics in (Medieval) Spanish have
a single form which may variously be construed as reflexive, direct or indirect object,
or as ethical datives (15).

(15) Testimonias *me* sed oy [Medieval Spanish]
witnesses me.dat be.imp today
‘Be witnesses on my behalf today.’

Yet another aspect of the clitics puzzle is that where there is more than one clitic in a
clausal string, they generally cluster together, so that for any statement purporting to
restrict the occurrence of the clitic pronoun to immediately following some preceding
category of expression or immediately preceding some verbal form, the statement has
to be complicated by the fact that another clitic may intervene between it and such a
host (data from Granberg 1988: 132):

(16) e ella dixo-*ge-lo* [...]  
and she.nom told.3sg-to-him.dat-it.acc
‘And she told it to him [...].’

(17) ca ya non te lo mandava matar  
because already not to-you.dat him.acc ordered.1sg kill.inf
‘because I no longer ordered you to kill him.’

Finally, there is the complication that the relative ordering of these clitics may vary
between closely related languages and even within a single language without any
distinction of interpretation. Of these, perhaps the most striking is French which
licenses pairs of third person clitics only in a DO-IO sequence (excluding the order *le
lui*), but requires pairs of third and first/second person clitics to occur only in the
inverse IO-DO sequence (as in the Spanish example (17)).4 The basis for such
clusterings is thus generally agreed not to have a semantic basis. But there is also
morphological idiosyncrasy, so that a purely phonological explanation doesn’t seem
appropriate either – in particular restrictions on Spanish clustering differ according as

---

3 The phrase *quid agis?* in Latin is generally used for ‘How are you?’ or ‘How’s it going?’.

4 We do not give French examples in detail, as a DS account of French remains to be developed.
the neutral dative form *se* is construed as ethical dative, indirect object, or reflexive. There is also notorious variation across dialects, with ‘leista’ effects in which the dative *le* appears to be spreading to include accusative uses, but with also ‘loista’ and even ‘laista’ dialects in which it is rather the masculine *lo* (or the feminine form *la*) which is becoming the form that can cover both direct and indirect object construals (Company 1998).

The intransigence of clitic positioning to syntactic, semantic, or phonological explication has led to debates as to whether these clusters interact with syntactic processes at all. In minimalist analyses, variant clitic properties are seen by some as associated with distinct features, hence distinct triggers for movement, inducing movement of the clitic to the requisite checking site (Cardinaletti 2008), others see them as subject only to feature-geometry forms of explanation (Cuervo 2005, Rezac 2008), yet others a mixture of the two (Adger and Harbour 2007). There have been debates over which clitics should have which features, what processes they trigger and, for those that argue for feature geometries, whether there should be rules making reference to concepts of domination displayed on the feature hierarchy (Heap 2005). In the majority of cases, the specifications proposed lack independent motivation, and so amount to little more than stipulated invocation of syntactic structure or feature geometry to directly reflect the idiosyncratic orders observed (see Chatzikyriakidis 2010 for detailed evaluation of this literature). Cardinaletti’s account of the array of idiosyncracies displayed in Italian (Cardinaletti 2008) involves distinguishing *gli* (the realisation of dative *le* when immediately preceding *lo*) as having a +person feature while its alternative realisation *le* has only a +number feature without that person feature. Rivero (2007), in addressing cluster properties in association with Spanish psych predicates, defines a newly distinctive mental-state +m feature, whose positing critically provides the necessary count of feature-strength to determine appropriate orderings on which her account depends. Adger and Harbour and others argue over whether there should be binary Participant, Author, and Hearer features over and above other features assigned, and there are debates as to whether Person should be posited as a feature at all (Anagnostopoulou 2005 “No”, Rivero 2007 “Yes” with both overt and covert variants) and over whether features should be binary (the Adger and Harbour 2007 account of the problematic morphological gaps posits both binary and non-binary features). Cuervo (2005) defines template positions onto which feature complexes have to be mapped (eschewing a movement-based account), and though noting the problems raised by morphological gaps, provides no account of them. Against these, structural accounts persist: Ormabazal and Romero (2007) for example argue for an agreement-based account, that for any language displaying VP-internal agreement, no more than one such agreement pairing is possible.

Things are little better in other theoretical frameworks. In optimality theoretic frameworks, for example, the set of constraints defined is highly particular to particular clusterings involving for example PERSONRIGHT, PERSONLEFT, EDGEMOST(Dat), EDGEMOST(Acc), and PARSE constraints (Grimshaw 2001, Legendre 2003), all defined to allow appropriate flexibility under appropriate conditions; but with the consequence that there is no restriction on possible clusterings, the constraints doing no more than matching the facts. In Monachesi (2005) and Anderson (2005), such clustering is taken to motivate the postulation of a morphology component defined as independent of either syntax or semantics, a move which means that lack of independent explanation of the data is turned into a design
feature of the grammar. Licensed co-occurrences per language are defined as varying morphological templates onto which the language-sequences have to be mapped, with no attempt in that system to explain why such clustering behaviour should occur. The overall impression from this increasing wealth of literature devoted to clitics is that there is little indication of anything approaching a principled explanation.

2.2 Word order variation

There is an unexpected twist on clitic variability which in this paper we wish to bring out. Looked at as a set of distributions, the sequence of clitics with respect to the verb is redolent of the patterning of full NPs relative to the accompanying verb in Latin, occurring regularly before the verb, and restrictedly after it. We shall argue that far from being a trivial observation, this is indeed the source of the explanation.

Latin constituent order variation is syntactically free in simple clauses at least, with NPs able to occur in any order and with any one or more NP able to occur before the verb, or after it. In consequence, there is no apparent indication from the order itself as to how the various parts are to be semantically combined:

(18) Catullus Lesbiam amavit [Latin]
    Catullus.sg.nom Lesbia.sg.acc loved.3sg.perf
    ‘Catullus loved Lesbia.’

Lesbiam Catullus amavit
Amavit Catullus Lesbiam
Amavit Lesbiam Catullus
Lesbiam amavit Catullus
Catullus amavit Lesbiam

It is, of course, the case specifications of the NPs which are largely responsible for the construal of the argument roles they project relative to the verb, rather than anything intrinsic to the ordering in the string. And it is these case specifications that get lost in the Romance systems. So it should be no surprise to find parallelism between Medieval Spanish and Latin distributions in the only set of nominals – the clitic pronouns – that retain some aspect of the Latin case system which was otherwise entirely lost. The assumption that case determines construal is however only partially true, in that, as with most case-marking systems, much of the Latin case system is syncretic, with only partial determinism of thematic role from the morphological form of the NP-expressions. Nominative and accusative forms of nouns are syncretic invariably in the neuter (as generally in Indo-European) and also regularly in the plural of the consonant stems. In the development of the Romance languages, phonological changes caused massive syncretism within nominal paradigms giving rise in Vulgar Latin to just two or three forms in many cases. For example, the first declension classical forms rosa, rosam, rosā, rosās, rosārum (singular nominative, accusative, ablative and plural accusative, genitive, respectively) are reduced to rosa while the late form rose stands for the rest of the paradigm, except for the dative/ablative plural. Ultimately, this led to a loss of case distinctiveness amongst the Romance languages (except for Romanian which retains oblique/non-oblique forms in certain declensions), and became general for NPs in Medieval Spanish. Syncretism also affected the weak pronominal system so that in Medieval Spanish some clitics, me and te amongst others, are not differentiated as to accusative/dative cases.
Despite such variable determinism in the case system in Latin, word order freedom extends beyond mere local “scrambling”, as constituents can be dislocated even across clausal boundaries:

(19) Stercilinum magnum stude ut habeas [Latin]
dunghill.sg.acc big.sg.acc ensure.imp.sg that have.2sg

‘See that you have a large dung hill’ (Cato De Re Rustica 6)

In these classic long-distance dependency constructions, case specifications cannot be seen as contributing anything more than a constraint on their construal, given their arbitrary dislocation from the expression on which they depend.

Nevertheless, despite such flexibility, word order in Latin is very far from being a total free-for-all. Even though more than one constituent can be dislocated and placed at the left periphery, in all cases involving dislocation from an embedded finite clause, there is invariably a restriction that all the constituents so dislocated must be interpreted as local to each other, as in (20, 21):

(20) [Ventus ad praefurnium caveto] ne
come-near.3sg.pres

‘Take care that the wind doesn’t blow on the furnace door.’ (Cato De Re Rustica 38)

(21) [digitum supra terram] facito semina emineant

‘Make the seeds project a finger above the earth.’ (Cato De Re Rustica 46)

This rigid local pairing of NP-expressions receives an echo in the subsequent clitic systems that emerged, with their rigid ordering before the verb, but essential locality with respect to each other. Until quite recently, surface word order had been taken to be a linearisation matter to be handled as a surface property not impinging on the structural core of syntax-internal mechanisms. But this leaves unexplained the rigid locality of any two such dislocated expressions relative to each other, a pattern that occurs quite generally with clitic sequences which, like multiple long-distance dependencies, cannot be split.

What we argue is that the patterns attributable to scrambling are indeed reflected in the distribution of Medieval Spanish clitics (see also Bouzouita 2008a, 2008b, 2011): and we will sketch an account that formally defines an explanation in these terms. In informal terms, local scrambling requires constructive use of case (Nordlinger 1998), with case specifications determining argument role in the presented structure in an online way in Latin. Long-distance scrambling, in which an expression can be dislocated arbitrarily far from its dependency site, indicates to the contrary that some case specifications do not perform any such local constructive role, but merely act as some kind of filter on appropriate identification of where they contribute to the overall structure. Multiple long-distance scrambling, in which pairs of such dislocated expressions may occur together at some early position in a string, can be modelled by
a mechanism that induces an essentially localised sub-structure, to be resolved in the overall structure as a unit. Finally, parenthetical construals can be available for any expression, so that some expressions can be analysed as in some sense independent of the structure within which they are contained. It is then the effects of these general mechanisms that underpin what has been seen as requiring clitic template specifications, with the various effects displayed in the distribution of clitics being modelled as a calcification of the sequences of actions which had in the earlier Latin system been induced by these general mechanisms, in different combinations.

An account of scrambling has been argued for in detail elsewhere with respect to Japanese and Korean (Cann et al. 2005, Kiaer 2007, Kempson and Kiaer 2010). Our primary aim in the first half of this paper is to show how that account as applied to Latin can be used as a basis for explaining clitic behaviours in Romance as the freezing of a set of parsing strategies, with the Person Case Constraint on clitic clustering effects explained as a result of the general tree-growth restriction preventing any of the offending clusters from ever emerging, notwithstanding the existence of apparently systematic exceptions.

3. Towards a Dynamic Syntax of Latin

The novel property of Dynamic Syntax as a syntactic theory is that the concept of structural underspecification and growth of interpretation intrinsic to processing is taken as the core syntactic notion. The syntax of the natural-language system is thus defined as a set of strategies for establishing the interpretation of some string of words in the order in which they appear, reflecting possibilities for choice in on-line parsing. The process involves the incremental development of tree structures representing a semantic interpretation for a string which are decorated by labels that progressively provide the information needed to determine the appropriate interpretation. Generation is defined in exactly the same terms: the very same rules apply in production as in parsing, the only difference between production and parsing being that whereas the parser may not know in advance the interpretation to be constructed, the producer in contrast must do so (Purver and Otsuka 2003, Purver et al. 2006). Hence, in generation there is from the outset a ‘goal tree’ which represents the interpretation to be conveyed, together with a defined constraint that in generation, each update step licensed by the parsing mechanism has to constitute an enrichment towards completing that ‘goal tree’ (formally a subsumption relation is required to hold between the parse tree and the goal tree; Purver and Otsuka 2003).

As the basis of the processing system is parsing, we begin by defining the general parsing strategies used in the framework. The starting point of this process is a tree with just a rootnode and a requirement to construct some propositional formula. The endpoint is a fully decorated binary branching tree structure encoding functor-argument structure of a familiar sort. As figure 1 displays, each completed interpretation is represented as a binary-branching tree whose rootnode is the propositional formula established and its daughter nodes the various sub-formulae that together yield this formula.

---

5 $\text{Fo}$ is a predicate that takes a logical formula as value, $\text{Ty}$ a predicate that takes logical types as values, $\text{Tn}$ a predicate that takes tree-node addresses as values, e.g. $\text{Tn}(0)$ being the rootnode. In general we omit the $\text{Fo}$ predicate in tree diagrams for simplicity. The $\hat{\circ}$ is a pointer, indicating the node currently under development.
The process of tree-growth is the sole basis of syntactic explanation: a sentence is defined to be well-formed just in case there is at least one possible route through that process that leads to a complete propositional tree with no requirements outstanding (see below). Tree growth involves the emergent unfolding of partial trees, whose node-relations and node-decorations all get progressively specified. Transition steps between partial trees are determined by a combination of general computational actions and lexical actions that are triggered by parsing words in the order in which they are presented in some string, together defining a monotonic process of tree-growth. These computational and lexical actions are expressed in exactly the same terms, that of growth along any of the dimensions associated with decorations on the trees defined by the system. Moreover, both sets of actions are defined using exactly the same vocabulary, allowing in principle for a sequence of computational actions to become associated with particular lexical items and subsequently stored as a lexically defined tree-update. The only essential differences between computational and lexical actions are that the former are, without exception, optional and not triggered by particular phonological (or orthographic) input, while the latter are so triggered and the actions they determine must be run.

Any aspect of tree construction or decoration may be partial. Accordingly, tree-relations, tree-decorations, type and formula specifications may all be only partially specified. Central to this process is the concept of requirement ?X for any decoration X representing a type, formula or treenode address. For example, decorations on nodes such as ?t, ?e, ?e → t etc. express requirements to construct formulae of the appropriate type on the nodes so decorated; ?∃.Fo(x) a requirement to provide a fixed formula specification; and ?∃.Tn(x)a requirement to provide a fixed treenode address. The underpinning formal system is a logic of finite trees (LOFT, Blackburn and Meyer-Viol 1994) with two basic modalities, <↓> and <↑>, such that <↓>α holds at a node if α holds at its daughter, and its inverse, <↑>α, holds at a node if α holds at its mother. Function and argument relations are distinguished by defining two types of daughter relation, <↓0> for argument daughters, <↓1> for functor daughters (with their

---

6 Quantification is expressed in terms of variable-binding term operators, so that quantifying NPs like all other NPs are of type e. The underlying logic is the epsilon calculus, the formal study of arbitrary names, with term-expressions whose internal structure is made up of an epsilon binder, ε, a variable, and a restrictor: e.g., ε, x, Mar(x). Since in Latin, nouns project full specification of terms, the structure defined to be projected by servum would be a subtree of which the quantifying term is the topnode, dominating a subtree decorated with binder, variable, and restrictor specification. We leave all details on one side.

7 This departs from the notation in Kempson et al. (2001) and Cann et al. (2005).
Inverses \( \langle \overrightarrow{0} \rangle \). Domination relations are then definable through Kleene star operators, e.g. \( \langle \overrightarrow{*} \rangle Tn(a) \) for some node identified as dominated by treenode \( Tn(a) \); a node decorated as \( \langle \overrightarrow{*} \rangle Tn(a), \exists x. Tn(x) \) is a node that is dominated by some node \( Tn(a) \) along some arbitrary sequence of mother relations. Such modal statements can be used to formulate requirements. These may be general, e.g. the requirements on an introduced proposition-requiring node for an argument-daughter node and a predicate-daughter node: such a node would be decorated with \( ?Ty(t), \langle \overrightarrow{0} \rangle Ty(e), \langle \overrightarrow{1} \rangle Ty(e \rightarrow t) \). Such requirements constitute subgoals on a wellformed derivation, and are filters on the output.

Requirements may also however be defined as lexically imposed filters on output: and this is the initial basis for modelling case specifications where this is structurally definable. For example, a nominatively marked expression is defined as projecting onto a subject node of the emergent tree an output filter requirement of the form \( \langle \overrightarrow{0} \rangle Ty(t) \) (the requirement that its immediately dominating node be of a formula of type \( t \)); an accusatively marked expression projects onto the immediate argument-daughter node of some emergent predicate-requiring node the requirement \( \langle \overrightarrow{0} \rangle Ty(e \rightarrow t) \). Thus case specifications, like all other generalisations, are expressed in terms of possible forms of tree growth. And so it is that a range of what in other frameworks are taken to be morphological or syntactic properties can in this framework be expressed as requirements on growth of semantic representation.8

Restrictions at the interface of syntax and semantics are also naturally expressible in these terms. An uncontroversial aspect of underspecification of content is that associated with anaphoric expressions, their intrinsic contribution to interpretation being that they provide only some partial specification of any occasion-specific interpretation, the particular value being determined by the context relative to which the uttered expression is understood. In this representational perspective, this is expressed by defining all such context-dependent expressions as projecting an interim place-holding device, adding to the basic Formula vocabulary the metavariables \( U, V, \ldots \), each associated with a requirement for a fixed value to be provided either from the context so far accrued in the interpretation process or subsequently from within the construction process. Whatever restrictions there are on the domain within which individual anaphoric expressions have to be construed are also defined in tree-growth terms as constraints on the (sub)-tree within which the values of metavariables have to be found. For example, in the case of reflexives, the value for the projected metavariable has to be found at some node \( Tn(a) \) along a path \( \langle \overrightarrow{0} \rangle \langle \overrightarrow{*} \rangle \langle \overrightarrow{0} \rangle Tn(a) \) from the node being decorated by the reflexive – that is from some co-argument along some unspecified but uninterrupted functor spine. Conversely, metavariables projected by pronouns cannot take such a local value, a constraint expressed as part of the process of substitution (see Kempson et al. 2001: 97).

---

8 The specification of case in these terms is naive in the sense that it assumes that particular cases determine directly the semantic function of the term projected by some noun phrase. This is not generally true (e.g. nominative expressions may be a semantic object in some passive constructions while other cases have ‘semantic’ counterparts). Some effects of this are noted below, but a more sophisticated theory of case in DS remains to be articulated.
3.1 Scrambling

More controversially, the very same perspective is adopted with what in other frameworks is taken to constitute evidence of either feature passing (Sag et al. 2002) or syntactic movement (e.g. Hornstein et al. 2005). Instead of positing morphologically empty sites in a string which are paired with some non-contiguous (left-peripheral) expression as a basis for articulating the contribution of that expression to interpretation of the string, a parsing-based perspective that follows the dynamics of processing of strings in real time presents such dislocated left-peripheral expressions at the early point in the string as associated with underspecified structural information that has to be updated later. Matching this explanation, one core mechanism is the license to construct a node dominated by some proposition-requiring node whose tree-relation is not fully specified with respect to that node. This is achieved by a rule of *Adjunction (read ‘star-adjunction’) which creates an “unfixed” node with precisely this property, described in the tree-logic language as $(↑*)Tn(a)$ with respect to some treenode $Tn(a)$. The exact role of such unfixed nodes is thus not specified at the point of introduction in the emergent tree structure, but is required to be determined at some later stage in the grammatical process by an enrichment process updating that underspecified relation to a fixed relation identifying its role as subject, direct-object, or indirect object.

There is also a more locally restricted process of introducing unfixed nodes (Local*Adjunction), for which an argument-node is constructed that is also underspecified with respect to some type-t-requiring node but with a tighter constraint that this relation be within the subtree from the node from which the underspecified tree relation is constructed. This is characterised on its introduction as having a modality $(↑0)(↑1*)Tn(a)$ with respect to some treenode $Tn(a)$. This specifies that the unfixed node is an argument $(↑0)$ that is related to an unspecified series of functor nodes to the dominating node $(↑1*)$. This has the effect of ensuring strict locality within a single predicate-argument array. Both underspecified tree relations are twinned with a requirement for update ($?∃x.Tn(x)$) so that a subsequent fixed tree-node relation must be provided in all wellformed derivations.

There are limits on how such underspecified relations can be constructed. A defining property of trees and the nodes they contain is that a node in a tree is uniquely defined by its relation to all other nodes in the containing tree (Blackburn and Meyer-Viol 1994). This has a consequence for the tree construction process that there can only be one unfixed node of a type at a time in any partial tree, as all such nodes are characterizable only by their relatively weak modality. This is not a constraint that has to be externally imposed: any duplication of some tree relation simply induces the immediate collapse of any such putative pair of nodes, which invariably leads to an incoherent treenode decoration unless the individual decorations of the duplicated nodes are compatible. It is this constraint, as we shall see in due course, that underpins Latin word order effects, restrictions on clitic cluster combinations in Romance and also Bantu object marker restrictions.

A common basis for cross-linguistic variation is the minor variation that lexical actions for related categories of expression in the differing languages may display. For example, with its relatively free word order and possibility of pro-drop, the parsing of a Latin verb induces a propositional structure whose argument nodes are decorated with metavariables, $U_{3sg}$, $V$...$, capturing the effect of null pronouns in such languages
without the assumption that these exist as parts of a linguistic string. The lexical entry in (22) illustrates the actions to be carried out by a parse of *amavit* ‘loved’, with the resulting partial tree shown in Figure 2 below.  

(22)  

\[
\begin{align*}
\text{IF} & \quad ?Ty(t) \\
\text{THEN} & \quad \text{put}(Tns(PAST)); \\
& \quad \text{make}(\downarrow_0); \text{go}(\downarrow_0); \\
& \quad \text{put}(Ty(e), \text{Fo}(U_{3sg}), ?\exists x. \text{Fo}(x)); \\
& \quad \text{go}(\uparrow_0); \text{make}(\downarrow_1); \text{go}(\downarrow_1); \\
& \quad \text{put}(?Ty(e \rightarrow t)); \\
& \quad \text{make}(\downarrow_1); \text{go}(\downarrow_1); \\
& \quad \text{put}(\text{Fo}(Amare'), Ty(e \rightarrow (e \rightarrow t)), [\downarrow] \bot); \\
& \quad \text{go}(\uparrow_1); \text{make}(\downarrow_0); \text{go}(\downarrow_0); \\
& \quad \text{put}(\text{Fo}(V), Ty(e), ?\exists x. \text{Fo}(x)) \\
\text{ELSE} & \quad \text{Abort}
\end{align*}
\]

This property is not shared by verbs in non-pro-drop languages whose argument nodes, as projected from the verb, bear the weaker characterisation of the requirement \(?Ty(e)\), without metavariables, thereby imposing the requirement of morphologically explicit argument expressions.

Given the common language of lexical and computational actions, the lexical projection of propositional structure by a verb freely interacts with the construction of nodes, for example by application of Local*Adjunction prior to the parse of a verb, where scrambling effects are driven by constructive use of case.

(23) Lesbiam Catullus amavit
Lesbia.acc Catullus.nom love.3sg.perf
‘Catullus loved Lesbia.’

Despite case specifications being defined to ensure that the term projected by some nominal expression is fixed in an appropriate position as an output filter, nothing

---

9 The applicability of specific rules or lexical actions depends on appropriate positioning of the pointer, \(\Diamond\), and while there is considerable freedom of the pointer back down a tree in anticipation of further development of nodes, movement of the pointer up the tree is highly restricted, and possible only if the type-requence on some node has been satisfied, and then, only to the immediate mother node or, in the case of unfixed nodes, to the node from which the underspecified relation was constructed.
prevents early update of the underspecified tree relation, by constructive use of case that guarantees the ultimate satisfaction of the output filter. The succession of steps required for the processing of (23) begins with the parsing of the accusative noun *Lesbiam* as decorating a locally unfixed node, followed by a series of steps of abduction to ensure that the accusative case constraint will be satisfied during the parse. Abduction proceeds in two steps: from the case constraint \(?\langle\uparrow_0\rangle Ty(e→t)\) on the argument-node to an annotation on the mother \(?Ty(e→t)\), to ensure that the accusative requirement is satisfied; and then from \(\langle\uparrow_1^*\rangle Tn(0)\) to \(\langle\uparrow_1\rangle Tn(0)\) to satisfy the tree-node requirement on the functor node.\textsuperscript{10}

A second step of Local*Adjunction takes place, and the parsing of *Catullus* is then taken to fix the value of the underspecified tree relation \(\langle\uparrow_0\rangle\langle\uparrow_1^*\rangle Tn(0)\) of Local*Adjunction into \(\langle\uparrow_0\rangle Tn(0)\) providing the basis for satisfying the nominative-induced requirement \(?\langle\uparrow_0\rangle Ty(t)\):

The result is that the relation between the argument node and the dominating node is fixed at the point of parsing the noun phrase, possibly well before the verb is processed. The output-filter restrictions of case-specifications serve thus to induce the update of an unfixed node to a fixed relation as each unfixed node is introduced. The actions of the verb then serve to fill out the remainder of the propositional structure to yield the appropriate output tree. These lexical actions operate exactly as before,

\textsuperscript{10}This sequence of steps can apply to all argument relations including subject: the Kleene* intrinsic to defining \(\langle\uparrow_0\rangle\) and other operators is satisfied by the empty set, so \(\langle\uparrow_0\rangle\langle\uparrow_1^*\rangle Tn(a)\) is true also of the subject relation.
giving rise to a duplication in the description of the tree of both subject and object nodes with the already constructed nodes being matched with nodes decorated by metavariables. This duplication of nodes harmlessly collapses into a single description for each affected node because metavariables are not part of the object language of formulae, but merely place-holders for such formulae. Therefore, the effect of the nodes constructed from parsing the two initial noun phrases is to provide the values for the metavariables projected by the verb.

The restriction that there can be only one unfixed node at a time remains satisfied, despite the application of procedures to build these nodes twice over. Nothing dictates which of these argument expressions is placed first, so the sequence of actions involving Local*Adjunction followed by a tree-update process reflecting the particular case specification can occur in any order, reflecting the freedom of constituent order which Latin displays. Given the restriction to only one unfixed node of a type at a time, this type of derivation is available only upon the assumption that on-line update of the tree relation is available, so no particular fixing of rule-order application is required: all other derivations will be precluded. And so it is that successful derivations to yield an interpretation of examples such as (23) can be built up incrementally.

This is by no means the only type of tree-growth sequence however. The first expression Lesbiam might be taken to decorate an unfixed node introduced through the non-local step of *Adjunction. In this case, by assumption, the case specification serves merely as a filter on update that is not immediately enriched to a fixed position, and in consequence no other unfixed node can be introduced by this step. As a discrete operation, Local*Adjunction nevertheless remains available for the processing of some matrix subject NP that might follow (Catullus in (23)). The consequence is that the sequence of strategies for constructing a string-interpretation pairing is by no means unique. Indeed arguably the only major difference in the way *Adjunction and Local*Adjunction apply lies in the fact that immediate case-update to a fixed tree relation cannot take place in theformer, because there is no presumption that the term is local to the primary predicate-argument array.

Unlike this alternative derivation of (23), a derivation involving *Adjunction is of course needed essentially for dependencies that are not local as in (19).

(19) Stercilinum magnum stude ut habeas
    dunghill.sg.acc big.acc ensure.imp.sg that have.2sg.pres
    ‘See that you have a large dung hill’

Furthermore, this similarity of processes underpinning long-distance and short-distance scrambling effects provides an immediate explanation for multiple long-distance dependency effects. With both processes involving the building of an unfixed node, we expect the possibility of a feeding relation between *Adjunction and Local*Adjunction, resulting in multiple long-distance dependency as in (21):

---

11 Equally, such NPs could be placed after the verb, since both for the application of the actions triggered by the verb and for applicability of Local*Adjunction, the pointer needs to be at the type-t-requiring node. We leave all details about post-verbal clitic placement for another occasion. In rigid verb-final languages, it is the details of tense-specification which ensure finality of the verb (see Cann et al. 2005, Kempson and Kiaer 2010).
(21) [digitum supra terram] facito semina emineant

finger.sg.acc above earth.sg.acc make.imp seeds.pl.nom/acc project.3pl

‘Make the seeds project a finger above the earth.’

In the DS account, such patterns are directly expected. *Adjunction would allow the construction of a propositional unfixed node decorated with the requirement ?Ty(t). Within in this unfixed propositional domain successive steps of Local*Adjunction may apply to construct partial propositional structures to be associated with the first two constituents in (21), on the assumption that prepositions can play the role of case in determining additional arguments for the accompanying predicate, yielding partial trees such as figure 5.\(^{12}\) In this way a sequence of argument nodes can be constructed. The position of the cluster of argument nodes is then resolved at a subsequent point in the construction process, in (21) with introduction of the propositional complement argument of *facito to yield the tree in figure 6.

\[\text{Figure 5: Parsing } \text{Digitum (supra) terram.}\]

\(^{12}\) On this analysis, we assume an account of prepositional phrases following Marten (2002), in which prepositional phrases function as optional arguments, and here we simply stipulate the relation of *supra terram as being that of third argument for some upcoming predicate. Many details of the analysis are omitted here, including the way in which the effect of the preposition is to over-ride the otherwise default construal of accusative as the highest argument within the predicate structure. The essence of the analysis stands, however.
What is notable about such intermediate structures in the present connection is the
construction of proposition-requiring structures which, at some intermediate juncture,
may contain only an array of argument nodes, as yet lacking the predicate node which
is essential to completing that structure.

There is one further general tree-construction strategy yet remaining before we have
anything approximating to a complete sketch of the mechanisms which the DS
framework licenses. There are also mechanisms for building paired structures, where
structures are taken to be twinned by being the result of a construction process which
ensures the sharing of some term in two such so-called linked trees. This process is
defined in DS for construal of relative clauses, clausal adverbials, and also external
topic constructions. Such secondary structures may have an attendant requirement that
the newly introduced proposition-requiring tree have somewhere within it a copy of
that term (specified as \(?\downarrow\)) \(\langle\downarrow\rangle\) \(\text{Fo}(\alpha)\): see Cann et al. (2005) for details).\(^{13}\)

The significance of this process for the overall DS perspective is that it extends the
range of alternatives whereby strings can be processed, so leading to an additional
possible structure as a form of interpretation, without the string itself necessarily
displaying an overt reflex of this additional alternative. For example, in a pro-drop
language, such a linked structure may indeed be decorated with a term provided by a
full NP, with the requirement that it be identified with one of the arguments of the
subsequent structure. They can be satisfied by information provided by the verb,
hence without need of any morphologically explicit anaphoric device. And should the

\(^{13}\) The process of inducing such pairs of semantic trees is permitted by defining an additional modal
operator in the tree logic \(\langle L\rangle\), and its inverse \(\langle L^{-1}\rangle\); and a rule is defined to yield a transition from an
arbitrary node in one tree across a Link relation to the toplevel of a new tree of whatever type.
NP taken to decorate the linked structure be a dative clitic pronoun, it may constitute some additional add-on to the remainder of the clausal sequence, without any further duplication of the information that it provides. Consider how the analysis of (14) (repeated here) might proceed.

(14) quid **mihi** Celsus agit?
    what me.dat Celsus.sg.nom do.3sg.pres
    ‘How, pray, is Celsus?’ (Lit. ‘What to me Celsus does?’)

Parsing the interrogative *quid* proceeds via *Adjunction to give an unfixed node and then, since all computational actions are optional, one move could be to construct a node linked to the main propositional node with the requirement to construct a term (?Ty(e)). The dative pronoun is parsed and the node is decorated with the name of the speaker, here assumed to be Horace (figure 8).\(^\text{14}\)

![Figure 8: Parsing *quid mihi*.](image)

By assumption, in this context, the term projected by *mihi* and identified as picking out the speaker, here assumed to be Horace, is not in this linked structure associated with a case constraint to find a particular function for the term so constructed (a polysemy effect which we shall see persists in Spanish); and the parse of the main clause continues. We end up with a tree like that in figure 9 indicating that the speaker, Horace, is only tangentially associated with the event denoted by the main verb, allowing, through normal inference driven by relevance considerations, a broad range of relations to be construed between Horace and what he has said.

\(^{14}\) In this analysis, no term is shared between the linked structure and the main proposition, making it like an analysis of gapless topics in languages like Chinese (Wu 2005).
The consequence of this flexibility is that there are a number of moves available at any stage of a parse sequence, in particular in the early stages when so little structural specification is as yet determined.

3.2 Processing pressures, word order and pragmatics

We now turn to how performance considerations might dictate preferred choices amongst these alternative strategies. General constraints on production and parsing will ensure that speakers and hearers maximise the use of context to cut down the need to search the lexicon for words expressing appropriate meanings or to employ inference to determine what is being conveyed. In particular, according to Relevance Theory (Sperber and Wilson 1995), with its trade-off between cognitive effect and expenditure of effort, pragmatic processes of utterance interpretation will tend to encourage the appearance of given material early on in a clause. Such positioning provides a means of minimizing the search within a given context to establish construal of pronominals as early as possible. This is of course no more than a relevance-based explanation of this well known given-before-new ordering.

However, to see the link between scrambling effects and clitic template restrictions, there is yet more to be said. In particular, different uses of pronouns in Latin developed into discrete encodings in the subsequent Romance languages. In the earlier Latin system, pronouns, like other nominal constituents, could be used either to provide some initial term which constitutes a point of departure for what follows, or to provide a contrast, an update to what follows. From a DS perspective, the first such effect would constitute the projection of a pair of independent linked structures, the second structure to be developed relative to the context provided by the first with a requirement of a shared term in the tree to be constructed. The second type of construal would involve the construction of an unfixed node by *Adjunction, decorating this with the term indicated by the initially placed expression, with anticipation of delay in updating this initially constructed node (see Cann et al. 2005). Both such devices are non-canonical in projecting structure that is not definitively local, and hence are characteristically associated with stress or distinguishing intonation as a signal that some non-canonical form of construal is required.15 Of

---

15 It has been argued in detail by Kiaer (2007) in connection with Korean that distinctive intonation is an important determinant of appropriate strategies for build up of the intended form of interpretation, in
course, we have no direct evidence of stress or prosody for a language such as Latin, but at the very least such contemporary evidence is suggestive.\footnote{The pronouns noted in (24) are taken by Adams (1994:104) to be illustrative of an emphatic use “often marked by placement of the pronoun at the head of its clause”.
}

(24) A: \textit{Tibi ego dem?}  
\hspace{1cm} 2sg.dat 1sg.nom give.1sg.sbj.pres  
B: \textit{Mihi hercle uero}  
\hspace{1cm} 1sg.dat by.Hercules in.truth  
‘A: Am I to give it to YOU?’  
‘B: Yes, by god, to ME.’ (Plautus \textit{Pseudolus} 626)

(24), for example, could be analysed in DS terms as being associated with the construction of a node introduced by *Adjunction to be decorated by a term representing the hearer.\footnote{The verb is omitted because the predicate, like the structural patterning, is recoverable from the context.} Such a device induces actions that by definition mark an emergent propositional boundary, being associated with introduction of a proposition-requiring node (decorated with \textit{?Ty(t)}) without any decorations other than the imposition of such a requirement. If, in anticipation of explaining the split that occurred between stressed and unstressed uses of pronouns, we turn to what the non-stressed uses of pronouns have in common, it is simply that they will lack this property: they will not be associated with those very structural devices which serve to identify some initiation of an emergent propositional structure, they will solely have a regular anaphoric function of context dependence. An interesting example of this occurs in (25) in which a strong pronoun (\textit{ego}), appears immediately before the two weak pronouns (\textit{te, ei}):

(25) \textit{quod scribis de illo Preciano iure consulto,}  
\hspace{1cm} what write.2sg.fut about that.sg.abl Precianus.sg.abl jurist.sg.abl  
\textit{ego te ei non desino commendare}  
\hspace{1cm} 1sg.nom 2sg.acc 3sg.dat not abandon.1sg.pres commend.inf  
‘Whatever you write about that jurist, Precianus, I do not stop recommending you to him.’ (Cicero \textit{Ad Familiares} 7.8.2)

The strong pronoun \textit{ego}, by analysis, decorates an unfixed node as the initial step in constructing some novel propositional structure following on from the building of an adjunct linked structure, and this choice clearly reflects a clearly emphatic form of construal. The actions of the weak pronouns that follow are then part of the progressive construction of this introduced structure, building, decorating and updating locally unfixed nodes.

Strong pronouns are of course not the sole means of introducing novel propositional domains. Other linguistic indicators of the emergence of a propositional structure include focused noun phrases, expressions containing a negative element, relative pronouns (26), complementizers (27), subordinate temporal adverbials, and verbs (28): indeed this is the only property common to this structurally heterogeneous set (examples culled from Adams 1994). Like their “strong” counterparts, positioning of particular to signal long-distance dependency effects, that depend on departing from the canonical build up of information locally.
pronouns under this use will be driven by relevance considerations for these, by assumption, are ever-present. This provides the functionalist underpinnings that explain the weak pronoun usage:

(26) quae \textit{tibi} nulla debetur
which.neut-pl 2sg.dat no.neut-pl is-owed
‘Nothing of which is owed to you.’ (Cicero \textit{Ad Atticum} 1.16)

(27) rogo \textit{ut} mi mittas dalabram
ask.1sg.pres that 1sg.dat send.2sg mattock.acc.sg
‘I ask that you send to me a mattock.’ (Terentianus 251.27)

(28) delectarunt \textit{me} tuae litterae
delighted.3pl.pres 1sg.acc your letter.nom.pl
‘I was delighted with your letters.’ (Cicero \textit{Ad Familiares} 9.16.1)

What these share is the characteristic that, once an emergent propositional structure is identified by some other expression, they will get placed as closely following as possible, decorating some locally unfixed node duly updated through its case specification, and so, like the strong pronouns, hugging the left edge of any such emergent structure as closely as commensurate with them not constituting a stressed/contrastive use.

4. Latin to Medieval Spanish

We now have everything in place to sketch out the assumptions a parsing perspective on grammar formalisms would lead us to expect in the explanation of the emergence of the clitic systems of the Romance languages from Latin. Medieval Spanish contains a codification of what had become two phonologically and functionally discrete uses of earlier pronominal forms: strong and clitic. What the clitic pronouns display is two distinct types of property: where they occur in a string, and what kind of tree update the clitic induces. On the one hand, since they constitute the only remaining reflex of earlier nominal case-marking, it is their triggers that are a direct reflex of the earlier set of environments that yielded pragmatic identification of propositional boundary marking, now encoding this information directly as calcified reflexes of that earlier more liberal system. Their positioning is like that of weak pronouns in Latin, i.e. following focussed elements, negative elements, complementizers, relative pronoun subordinators and verbs (for a detailed account see Bouzouita (2008a) from which these data are taken):

(29) Esto es el pan de Dios que \textit{vos} da a comer
this be.3sg the bread of God that you.dat give.3sg to eat.inf
‘This is the bread of God that he gives you to eat.’

\footnote{Examples of the other types of left-edge identifiers can be found in Bouzouita (2008a, 2008b, 2011) and Cann and Kempson (2008).
\footnote{Following Sperber and Wilson (1995), if there are specific inferential effects to justify commensurate enlargement of the context to be searched, this would explain the lack of tight correspondence between weak pronoun positioning in Latin and any fixed second position noted by Adams (1994), even assuming that such putative second-position effects are clause by clause (or “colon” by “colon” to use his terminology).}
(30) e dize que lo tenie del prior de Sancti Johannis and said.3sg that it.acc had.3sg of-the prior of Saint John
‘And he said that he got it from the prior of Saint John.’

(31) Connocio-la Jacob
recognised.3sg-her.acc Jacob
‘Jacob recognised her.’

In Latin, as we have already seen, a sequence of NPs (Devine and Stephens 2006) can cluster before the verb, and this pattern too emerges in Medieval Spanish with the clitic pronouns:

(32) Et los dioses me quisieron mal e me lo quieren
and the gods me.dat wanted.3pl harm and me.dat it.acc want.3pl
‘And the gods wanted to harm me and they (still) want to.’

4.1 Placement of clitic pronouns: the production pressures
On the other hand, the structural relations induced by the clitic pronouns also show similarity to NP distribution in Latin, and this is not just a trivial continuation of fixed argument relations associated with individual clitics, for though some induce a fixed node for a given argument relation, others induce an underspecified node without fixing the argument relation, and yet others induce pairs of nodes. Rather, the range of update actions provided by the clitics matches the variation in update actions which a sequence of computational actions plus lexical actions provided in the earlier Latin system. First, there is the building of a fixed tree relation. The accusative clitic displays a fixed interpretation corresponding to the construction of a fixed structural relation, with the non-syncretic accusative forms, lo, los, and their feminine-marked counterparts signalling only direct object function (data from Granberg 1988: 135):

(33) Al senor lo faras
to.the gentleman it.acc you-will-do
‘You will do it to/for the gentleman.’

(34) cuando lo ganó
when it.acc won.3sg
‘When did he win it?’

These echo the earlier free availability of Local*Adjunction plus abductive update, here apparently lexicalised into a macro of actions leading directly to a fixed tree relation (see also Bouzouita 2008a, 2008b). Given that the only difference between computational and lexical actions may be that the actions in question become associated with a lexical trigger, the construction of a fixed tree relation is only one such possibility. There is also the action of constructing a locally unfixed node as though by Local*Adjunction. And, in this connection, the dative clitic and the first and second person clitics arguably induce the construction of an underspecified structural relation, displaying, as they do, a large range of interpretations. The consequence of this lack of determination of interpretation is that their contribution to
the emergent structure may not be able to be determined immediately, but only in combination with the verb with which they are associated:\textsuperscript{20}

\begin{flushleft}
(35) Yo \textbf{vos} defiendo que non vengades y mas et si non yo \textbf{vos} cegaré et \textbf{vos} mataré
\end{flushleft}

‘I forbid you to come and if not, I myself will blind you and kill you.’

In (35), the first occurrence of \textit{vos} is construed as indirect object, the second and third as direct object; but the morphological input is undifferentiated between these, suggesting that these reflect the construction of a locally unfixed node without update, leaving the relatively weak structural relation having to be updated by the later projection of structure by the verb. And finally there are the clitic clusters, which occur in the same relative position as the singleton occurrences, sometimes written as a single item, inducible as an individualised lexical sequence of actions reflecting the earlier building of construals of clustered NPs by a combination of *Adjunction and sequenced combinations of Local*Adjunction plus update, all listed as a single lexical entry with actions to induce the construction of a cluster of argument terms/nodes.

The range of effects we see displayed in the clitic pronouns of Medieval Spanish is thus broadly the range of effects seen in local scrambling in Latin. This is precisely what we would expect in a transition in which the availability of case specifications on a general basis disappears, being replaced by case specifications only within the pronominal system. As noted above, in the DS framework, general computational actions and lexical actions are expressed in exactly the same terms. Lexical actions, like their general counterpart, characteristically induce the construction of nodes in some partial tree in addition to providing decorations for the nodes which the actions associated with the word in question trigger. Thus a shift in tree-update actions from a sequence of general actions inducing nodes for which words provide decoration to a macro of actions associated with an individual word inducing both structure and decorations is exactly what one might expect in a shift from general to lexically triggered actions. And in this shift, any one word would normally be associated with only one such sequence of actions (unless its precursor in the source language was ambiguous): and so it is that the various clitic pronouns reflect one or other such action-sequence. Seen in processing terms, the clitic-template phenomenon is thus a freezing of scrambling strategies, hence explicable as a progressive shift, each lexical specification reflecting one of a set of strategies for early NP placement.

\textit{4.2 The Person Case Constraint explained}

But we can go further than this, as we now have a ready explanation of the Person Case Constraint. This, recall, was the non-cooccurrence of first and second person clitics, and the non-cooccurrence also of first/second person clitics with a third-person dative clitic. Both variants of the Person Case Constraint now fall into place: the morphological gaps follow from the tree-logic restriction that there can be no more than one underspecified tree relation of a type at any point in the tree-growth process. The power of this explanation is that it automatically provides an explanation for why

\textsuperscript{20} Notice in (35) the initial strong pronoun \textit{yo} ‘I’, in contra-distinction to the weak form \textit{ho} in the following conjunct.
the gaps in the clitic template possibilities associated with the Person Case Constraint do not occur.

4.2.1 The Strong Person Case Constraint

Let us take the more comprehensively satisfied restriction first, the preclusion of any co-occurrence of first or second case specifications construed as DO with a third person dative specification construed as IO (the so-called strong form of the Person Case Constraint taken to hold whether or not the form is syncretic: Bonet (1995), Nevins (2007), Ormabazal and Romero (2007)). Recall that there was no need of stipulation that there should be only one unfixed node of a type at a time: in all putative cases where more than one such underspecified tree-relation might be introduced, they collapse as undifferentiatable, with all cases where the resulting treenode decoration is inconsistent being necessarily debarred. This is precisely the scenario which these morphological gaps present. Given the analysis of dative as intrinsically underspecified as to whether the node being decorated is a direct or indirect object (or a semantically weak adjunct), the syncretic first and second person forms will be predicted not to co-occur with any such form, irrespective of order, since they too have a form that fails to discriminate between the various argument roles they can satisfy. Upon an analysis of tree growth that reflects this underspecification, both must be taken to decorate a locally unfixed node. Neither 1st or 2nd person markers could accordingly ever be constructed together with a third person dative marker, let alone be constructed sufficiently often to get routinized into a stored clustered form: both are defined as inducing the construction of a locally unfixed node without any case basis for inducing appropriate update ahead of the verb. Their lack of co-occurrence is immediately predicted. It is not the occurrence of these syncretic forms construed as indirect object with an accusative third person form which is problematic. Indeed, it is not the specific construals of these pronouns that provide the appropriate explanation for the oddity of the precluded forms. It is the fact that these forms, being syncretic, are associated with inducing only the building and decorating of some locally unfixed argument-relation, and so cannot co-occur with a dative or any other case-specification which is itself associated with inducing exactly the same weak tree relation.

It might be argued that this falls into the trap of identifying case underspecification with structural underspecification, equating gaps in a paradigm with syncretism. As pointed out by Adger and Harbour (2007), accounts which turn on case syncretism as reflecting relative weakness of specification are at best insufficient, since the same restriction is displayed in clitic systems with no syncretism in the clitic forms. In particular, this is displayed by Greek, with its distinct nominative/genitive forms for both first/second person subject and object marking:21

(36) *su me sistisan
2sg.gen 1sg.acc recommended.3pl
‘They recommended me to you.’

Yet, as it turns out, such examples buttress the DS account, for they illustrate the one further type of tree growth that the DS system leads us to expect. So far, we have

---

21 In Greek, the indirect object relation is expressed by genitive case. For a detailed account of PCC effects, see Chatzikyriakidis and Kempson (2011).
itemised the induction of an unfixed argument node that is taken to give rise to immediate enrichment, the induction of a linked structure, and the building of sequences of locally unfixed nodes from an intermediate node. But we haven’t had an instance analogous to the core mechanism underpinning long-distance dependency, which is the specification of case as decorating an unfixed node which does not induce immediate update. But this is the scenario provided by Greek. In these cases, the morphological specification for direct and indirect object arguments is distinct. If however, we assume that one of the options for tree growth that might get calcified is precisely such a non-constructive use of case, then we have the basis for analysing Greek, despite the lack of syncretism in first and second person clitic pronouns. All the clitic forms decorate an unfixed node, each with a filter on output encoding the appropriate constraint, with \(?(\uparrow_0)Ty(e \rightarrow t)\) for accusative, and some weaker specification for dative, dative being type-underspecified as between adjunct and argument constructions and so of necessity underspecified as to tree position. This account of Greek clitic case specifications as frozen reflexes of case filters has the added bonus of completing the picture of possible calcification updates that clitic systems might reflect. So one type of problem isn’t a problem for the analysis at all: to the contrary, it buttresses it.

4.2.2 The weak variant of the PCC
This explanation of the Person Case Constraint in terms of a structurally weak relation and no more than one unfixed node of the same type at a time should, without doubt, carry over to anticipate equally that co-occurrence of first and second person pronouns should also be impossible, at least in a language such as Spanish in which the forms are syncretic. In many languages first and second person clitics are indeed mutually complementary, and indeed, as we would expect on the analysis just outlined, they are precluded in Greek:

(37) *Mas se edosan  
    1pl.gen 2sg.acc give.3pl.past  
‘They gave you to us.’

(38) *Sas me edosan  
    2sg.gen 1sg.acc give.3pl.past  
‘They gave me to you.’

Surprisingly, however, many cases are fully acceptable in Modern Spanish:

(39) No te me acerques  
    not you.refl me.dat come-closer  
‘Don’t come closer to me.’

Examples like these might be taken to indicate that at least this subcase of the PCC is not grounded in a strong structural restriction, indeed is no more than a reflection of the fact that events describable by ditransitives in which both participants described are human are not common, and so didn’t happen to lead to routinisations and encodings in the clitic clusterings that emerged (Haspelmath 2004). However, there is evidence that the stronger structural explanation is correct, simply obscured by the presence of polysemy of the relevant dative for some languages, specifically in Spanish (and Latin). It is certainly the case that in all languages, dative construals
show a flexibility between an adjunct vs an argument form of construal, suggesting
the necessity of saying that the dative is intrinsically underspecified for its logical
type, and so must be associated with relatively weak structural specification, at least
upon one analysis. Furthermore, in many languages, the apparently adjunct form of
construal includes so-called ethical datives, in which the expression which is dative-
marked may be only loosely linked with the predicate associated with the verb,
characteristically associated with first and second person forms as utterance
participants, implied to be indirectly affected by the event described by the verb plus
its arguments. Spanish has very rich use of such datives, as did Latin before it, as in
the Medieval Spanish examples below:

(15) Testimonias **me** sed oy
      witnesses **me.dat** be.imp today
      ‘Be witnesses on my behalf today.’

(40) y **te me** devuelvan vivo
     and **you.acc me.dat** bring-back alive
     ‘and may he bring you back to me alive’ (‘may he bring you back alive for my
      benefit’)

(41) **Me le** gritaron a mi hijo
     **me.dat him.acc shout.3pl at my son**
     ‘They shouted at my son (and that affected me).’

Greek also has ethical datives:

(42) **mou arostise to pedi.**
     1sg.gen was-ill the child
     ‘The child was ill on me.’ (‘The child was sick and this concerned me.’)

The question is whether such dative construals are sufficiently distinct to warrant a
discrete lexical basis. Cuervo (2005), arguing for feature-geometry style of analysis,
provides extensive evidence that in Spanish they do. Such ethical dative construals
also occur in Greek, but the distribution of dative clitics is much more restricted than
in Spanish:

(43) **mu ton** malosan to gior mu
     1sg.gen 3sg.m.acc shouted-at the son mine
     ‘They shouted at my son (and that affected me).’

As already noted, first and second person clitics never co-occur in Greek; and, in
addition, no more than one dative clitic is ever possible (unlike the Medieval Spanish
double dative example in (41)). Finally, in Spanish, under certain circumstances, more
than two clitics may be possible, indeed sometimes with more than one dative clitic
(data from Cuervo 2005, though acceptability judgements are very variable).22

---

22 This is never possible in Greek, which has nothing analogous to these data.
If we follow up on the Cuervo (2005) analysis within a DS perspective, we find a natural basis for distinguishing Spanish (under these variants) and Greek. On the one hand, the ethical dative as an adjunct construal can be analysed as inducing an independent LINK relation, so that all such dative instances would be decorating a node within a distinct tree, quite unlike the analysis of the dative that within an individual structure induces an unspecified tree relation. On this assumption, Spanish is seen as having evolved homonymous dative forms, the one initiating a transition onto a linked structure, the other inducing a weakly specified tree relation within a single structure. On this analysis we expect wellformed examples such as (39), despite the preclusion by the system of two unfixed nodes of a type at a time, since for all the apparent PCC violations, an alternative strategy is available in which only one unfixed node is constructed, the other clitic being taken to induce the construction of a fixed LINK relation, hence not in conflict with the tree-structural restriction. More generally, since all the dative clitics, first, second and third person, all allow ethical dative construals, on this account, we expect all combinations to be wellformed, even though not perhaps occurring often enough to have become a stored, routinized pairing. Furthermore, as independently noted by van Hoecke (1996), ethical datives and argument-construal of datives merge seamlessly into one another, in particular for all first and second person clitic pronouns, since all first and second person specifications by definition constitute specification of the speech participants and their relation to the event described, so that there are grounds for positing an analysis in terms of a LINK transition for a dative pronoun without necessarily restricting the applicability of such a strategy to any particularly idiosyncratic non-argumental role.

---

23 A further form of explanation for the rare cases of co-occurring first and second person clitics that can be observed in Medieval Spanish, as indicated by the scribal transcription of the pair of clitics in (i):

(i) Qui-**d** nos dio por alcalde?
    who-you.acc us.dat gave.3sg as mayor?
    ‘Who gave you to us as mayor?’

Notice the phonological cliticization of the second person on the wh-form, suggesting these are the result of an early step of *Adjunction feeding the building of clustered subject and object argument nodes associated with that first unfixed node. Under this derivation, the subsequent occurrence of *nos will be able unproblematically to decorate a node locally unfixed with respect to the root, even though the ultimate position of the first cluster is itself not resolved until the verb is parsed.
On this analysis, the challenge is to explain why a language such as Spanish can sidestep the weaker variant of the PCC but not the stronger. This is straightforward. The stronger form of the PCC is a restriction on co-occurrence of first/second person pronouns construed as accusative and a third person dative construed as an indirect object, i.e. on how internal subcategorised arguments to a predicate can be realised by a clitic cluster. But with this limitation on the range of interpretations for both first/second person pronoun clitics and on the third person dative clitic, the type of construal for either of these clitics cannot be expressed through the alternative strategy of a LINK relation, for this is by definition external to the array of structure associated with a predicate plus its arguments. Hence the universality of the restriction, even applying in a language such as Spanish which allows me/te forms to co-occur.

This account of the potential for variability in the extent to which a language can sidestep the restriction underpinning the PCC resides in whether homonymy has emerged in the diachronic development of the language. This provides an immediate clue to the cross-language variation: we can see the languages in which both variants of the PCC hold as languages in which the ethical dative construal did not develop into any discrete homonymic form. In particular, one might argue, Greek presents such a case. In these more restricted languages, the dative clitic simply has the weakly specified form, inducing a structural relation that is compatible with a number of construals; but, with no alternative set of actions to induce, no first and second person clitics will ever co-occur, nor, more generally, will there ever be more than one dative clitic. So the language differentiation turns on whether, as in Spanish (at least in some dialects), the dative is polysemous between a characterization that induces a LINK transition in addition to the characterization that induces a weakly specified structural relations, or whether, as in Greek, the dative has merely a single specification. Hence the Person Case Constraint can be seen as grounded in a strong universally sustained restriction, with all apparent violations of the constraint explicable through the effects of variation in construal leading to the encoding of a distinct strategy. This account has the further advantage that one would expect exceptions only on a lexical basis, given the grounding of these exceptions in polysemy. With these set aside, the constraint holds absolutely: the gaps in the morphological paradigms in question arise because the individual sequence of actions to induce the precluded tree growth process could never have occurred, let alone have occurred often enough to have become routinised as a lexically triggered macro of actions.

5. The PCC in the larger perspective

With the PCC data seen as a mere consequence of a much more general structural constraint, the account can be evaluated by the potential applicability of this type of explanation. A structural constraint of this generality should be expected to have reflexes elsewhere in the grammar, acting as a constraint in quite different areas. Here we can do no more than signal the type of case one might expect, as indicative of the methodology of argument that this form of explanation leads to. Phenomena where, puzzlingly, only one type of relation is possible, despite apparent diversity in the structural effects, constitute a case in point. One such is a pattern in the verbal morphology of many Bantu languages, here illustrated with Otjiherero, spoken in Central and Northern Namibia and by a smaller group of speakers in Botswana. As in other Bantu languages, Otjiherero displays a complex prefixed sequence of subject, tense and object markings. These prefixes are very generally syncretic. For example, vé is a class two (human plural) marker which can be construed as either direct or
indirect object. These sequences are subject to some puzzling restrictions. One such restriction is in the object marking. Otjiherero agreement marking ranges freely over indirect object, direct object and locative-adjunct indications, yet despite the fact that their class-marking would serve to distinguish them, only one possible object marking is allowed, construed as either indirect or direct object or as locative:

(48) ú-térék-èr-à òvá-éndà ònyàmà p-òngàndà.
    sm1-cook-appl-fv 2-guests 9.meat 16-9.house
    ‘S/he cooks meat for the guests at home.’

(49) ú-vé-térék-èr-à ònyàmà p-òngàndà
    sm1-om2-cook-appl-fv 9.meat 16-9.house
    ‘S/he cooks them meat at home.’

(50) ú-i-térék-èr-à òvá-éndà p-òngàndà
    sm1-om9-cook-appl-fv 2-guests 16-9.house
    ‘S/he cooks it for the guests at the house.’

(51) ú-pé-térék-èr-à òvá-éndà ònyàmà.
    sm1-om16-cook-appl-fv 2-guests 9.meat
    ‘S/he cooks meat for the guests there.’

Trying to use two object markers leads to ungrammaticality:

(52) *ú-vé-i-térék-èr-à p-òngàndà
    sm1-om2-om9-cook-appl-fv 16-9.house
    ‘S/he cooks it (for) them at the house.’

(53) *ú-i-vé-térék-èr-à p-òngàndà
    sm1-om9-om2-cook-appl-fv 16-9.house
    ‘S/he cooks it (for) them at the house.’

There are thus grounds to warrant the hypothesis that this entirely distinct Bantu-internal morphological problem is subject to explanation in terms of the very same constraint as the Person Case Constraint. An analysis along these lines is developed in Marten et al. (2008), relating the restriction in Otjiherero of one object marker in the verbal cluster, and the absence of object marking in passives, to the same constraint of having only one unfixed node of the same type at a time. Furthermore, variation in object marking across different Bantu languages (see e.g. Marten and Kula 2012 for an overview) is in many respects reminiscent of variation encountered in Romance (e.g. Cocchi 2001, Labelle 2008), and thus might be explicable in similar terms to the analysis presented here. Another piece of evidence comes from the same language group but from a very different domain – in the Tanzanian Bantu language Rangi,
future tense is regularly marked by fronting of an infinitival verb form and a following
inflected auxiliary. In the analysis explored in Gibson (2013), this reflects the
development of an unfixed node which is decorated with information from the
infinitive, and which is fixed when information from the auxiliary provides fixed tree
structure, with the subject marker decorating a locally unfixed node. Interestingly,
when another element decorating an unfixed node is present in the clause – such as a
wh-expression, a negation marker or a focused NP – the infinitive follows the
auxiliary, even though the future tense interpretation is maintained. In DS terms, this
follows from the restriction on only one unfixed node at the time, explored in detail in
the preceding sections.

6. Conclusions
The explanation of the complexity of clitic clustering presented in this paper competes
on the one hand with accounts which have been taken to justify the specification of
morphological templates within some syntax-independent morphology component,
and on the other hand with accounts of such clusters in terms of feature-geometry
which are at best only a trigger for structural processes and are not intrinsic to the
structural processes themselves. Relative to these, the present perspective suggests a
much stronger and more restrictive alternative – that morphosyntactic phenomena,
like syntactic processes more generally, can be explained solely in terms of the
dynamics of the ongoing process of building up interpretation, with morphosyntactic
particularities displaying frozen reflexes of these general structure building processes.
Whether this novel perspective on morphosyntax and syntax can be sustained as a
general hypothesis might be seen as remaining an open question, but the
competitiveness of the present account of the PCC against current alternatives gives
grounds for optimism.

References
Latin. Transactions of the Philological Society 92, 103-178.
Adger, D. and D. Harbour. (2007) Syntax and syncretisms of the Person Case Constraint. Syntax 10, 2-
37.
607-647.
Cooper and R. Kempson (eds.) Language in Flux: Dialogue Coordination, Language Variation,
Kempson, E. Gregoromichelaki and C. Howes (eds.) The Dynamics of Lexical Interfaces, 89-130.
Stanford: CSLI Publications.
Bouzouita, M. and S. Chatzikyriakidis (2009) Clitics as calcified processing strategies. In M. Butt and
pronouns. In R. Cooper and R. Kempson (eds.) Language in Flux: Dialogue Coordination,
Language Variation, Change and Evolution, 179-220. London: College Publications.


Directionality and the Production of Ungrammatical Sentences

ASH ASUDEH
University of Oxford & Carleton University
ash.asudeh@ling-phil.ox.ac.uk

It is a puzzling fact that English speakers produce resumptive pronouns under certain circumstances, yet reject sentences containing resumptives as ungrammatical. This paper reviews a language production model that attempts to solve this problem through prioritization of local well-formedness over global well-formedness and, more specifically, considers the model’s implications for directionality in grammar. In particular, English resumptive in islands imply that the unbounded dependencies in grammar are filler-driven, featurally defined at the top of the dependency, and that islands are identifiable by the grammar from outside the island, in a top-down fashion. Lexical-Functional Grammar provides the formal framework for incrementality in the production model, which is based on monotonic information growth in LFG grammars. LFG also provides a formal, ‘outside-in’ theory of unbounded dependencies that treats them as filler-driven and allows island barriers to be identified from outside of the island, rather than from the inside.

1. Introduction

The language faculty is a knowledge system that characterizes linguistic competence. Derivations in syntax, as laid out most influentially in the work of Chomsky, are understood as a formal mechanism in the system that characterizes this knowledge; i.e. derivations are a component in the specification of the set of well-formed expressions in a language. As part of linguistic competence, derivations are not necessarily understood as specifying the order of steps in language processing, which is an aspect of linguistic performance. Students of transformational grammar are discouraged from interpreting derivational syntax as embodying claims about the order or nature of processing. In this context, the issue of whether linguistic structures in the competence system are computed top-down vs. bottom-up or left-to-right vs. right-to-left is a purely algorithmic aspect of how the competence system works. For non-derivational, constraint-based theories of syntax, such as Head-Driven Phrase Structure Grammar (HPSG; Pollard and Sag 1987, 1994, Ginzburg and Sag 2000) and Lexical-Functional Grammar (LFG; Kaplan and Bresnan 1982, Bresnan 2001, Dalrymple 2001), the question of directionality seems yet more immaterial. Linguistic structures in these theories must satisfy a set of unordered constraints, so the result is always equivalent, no matter the algorithmic starting point for solving the constraints.

Nevertheless, the competence system is embedded in a processing system and directionality is an important part of the latter system (Jackendoff 2007). Parsing/understanding works from form to meaning and production/generation from meaning to form. Moreover, it is a standard view in psycholinguistics that parsing and production are incremental. Increasingly, however, incrementality has become a locus of explanation of facts of the competence system itself. For example, Phillips (2003) proposes the “Incrementality Hypothesis”, which states that “Sentence structures are built incrementally from left to right”,

I would like to thank the following people for helpful comments, suggestions and discussion: Emily Bender, Joan Bresnan, Mary Dalrymple, Nicolas Guilliot, Philip Hofmeister, Florian Jaeger, Peter Sells, Neal Snider, Mike Tanenhaus, and audiences at Carleton University, the Hebrew University of Jerusalem, the Linguistics Association of Great Britain, the University of Canterbury, and the University of Rochester. Any remaining errors are my own. I would also like to thank Cristiano Chesi for patiently shepherding this paper. This research was supported by an Early Researcher Award from the Ontario Ministry of Research and Innovation and NSERC Discovery Grant #371969.

© Copyright 2013 Ash Asudeh
STIL – Studies in Linguistics. 6:83-106
Special Issue on “Directionality of Phrase Structure Building” ISSN: 2281-3128
as part of an explanation of conflicts in tests of syntactic constituency. Incrementality is a cornerstone of constraint-based syntactic theories, where it is an automatic consequence of the formal system. Incrementality is formally captured as monotonic information growth: generally, as the set of constraints that describe the linguistic structure grows larger, so does the structure that satisfies the constraints (Bresnan 2001: chapter 5).


I will argue that outside-in functional uncertainty together with a natural account of local well-formedness of fragments yields an explanation of otherwise puzzling data to do with resumptive uses of pronouns and other nominals in English, as in (1).

(1) *Mary knows the man who Thora saw him.

Most dialects of English do not have grammars that license resumptives (Chao and Sells 1983, Sells 1984), so sentences like (1) are generally judged as ill-formed. Sells (1984) introduced the term “intrusive pronoun” for these ungrammatical resumptive-like pronouns.

The puzzle is the following:

(2) **Intrusive Pronoun Puzzle:**
Native speakers reject intrusive pronouns as ungrammatical, yet produce them.

This behaviour is distinct from other superficially similar cases, such as center embedding (Chomsky and Miller 1963). In these other cases, a speaker rejects a form as ungrammatical but also would not be expected to produce it. The usual explanation of center embedding and other relevant phenomena is that a speaker fails to parse or produce, due to performance reasons, a sentence that is underlyingly grammatical. In Asudeh (2011a, 2012), I have proposed very much the opposite sort of explanation for intrusive pronouns: what speakers are doing with intrusive pronouns is producing a form that is underlying ungrammatical. The form is produced due to the processor’s prioritization of local well-formedness over global well-formedness.

In this paper, I consider the theoretical implications of this proposal in light of the phenomenon of intrusive pronouns in islands, as in (3).

(3) John would like to meet the linguist who Peter knows a psychologist that works with her.

If the explanation for the Intrusive Pronoun Puzzle (IPP) is that local well-formedness is prioritized over global well-formedness, then the explanation of intrusives in islands can be assimilated to the explanation of the IPP only if the island structure itself is locally well-formed. The island cannot be locally well-formed in a theory that defines islands from the inside-out, or in terms of some kind of notion of bounding node or barrier or phase that blocks upward movement of a constituent that originates in a position within

---

1 This statement is qualified for two reasons. First, mechanisms have been introduced that slightly relax the monotonicity requirement on a per constraint basis. The mechanism in HPSG is defaults (Lascarides and Copestake 1999). The mechanism in LFG is restriction (Kaplan and Wedekind 1993). Restriction is not strictly non-monotonic, because a subsumption relation is preserved between relevant structures, but addition of a constraint containing the restriction operator does not result in structural growth. Second, as will become clearer below, there are constraints in LFG that check structures but which do not themselves add information. Addition of these constraints to the set of constraints describing a structure does not result in structural growth, but a system containing the checking constraints but not restriction is strictly monotonic, because addition of constraints in the absence of restriction never results in a less specific structure. Moreover, if we construe monotonicity as a property of the set of constraints itself, rather than of the structures themselves, then even the system with restriction is strictly monotonic (Ron Kaplan, p.c.), because the equation containing the restriction operator is still an equation that is added to the set of constraints such that the set of constraints grows; similarly, the checking constraints also enlarge the set of constraints.
the island (Chomsky 1973, 1986, 2001), or in a theory that prevents island structures from being built at all (cf. Creswell 2002 on the problem of resumption for islands in Tree Adjoining Grammar). Therefore, it is reasonable to conclude based on intrusive pronoun data that, if a theory seeks to capture the relevant sorts of islands in the competence grammar, island constraints must be definable from outside the island, such that the island itself is locally well-formed. If island constraints are defined form outside the island, this in turn favours a theory of unbounded dependencies that involves top-down search in the grammar for the integration site of the constituent at the top of the unbounded dependency (e.g., a wh-pronoun). Thus, the explanation of the IPP in Asudeh (2011a, 2012) has a consequence with respect to directionality: the integration of unbounded dependencies in grammatical structure must be filler-driven not gap-driven.

2. The Problem

The fact that English speakers produce intrusive pronouns and other nominals in resumptive contexts has been known for quite some time. Kroch (1981: 129) discusses examples like the following, which contains an intrusive epithet:

(4) There was one prisoner that we didn’t understand why the guy was even in jail.

Kroch (1981) proposes an incremental model of speech production that crucially generates a filler before planning of the sentence has been completed. As production proceeds, the system enters a situation where the intended base position of the filler-gap dependency is in an island or would violate the that-trace filter/Empty Category Principle (ECP). An NP is inserted to avoid the violation, a kind of transderivational repair strategy. In short, Kroch’s proposal is that some nominal, typically a pronoun (but not necessarily, because of examples like the one above), is inserted to avoid a grammatical violation due to poor planning.

In subsequent work, Prince (1990) showed, based on a collection of naturally occurring utterances, that the insertion of intrusive nominals is quite pervasive. Examples such as the following illustrate that intrusive pronouns occur even in situations where a gap would actually be well-formed (Prince 1990: (4c) & (15d)):

(5) I have a friend who she does all the platters.

(6) You get a rack that the bike will sit on it.

Examples like these cast serious doubt on the viability of intrusives as a repair strategy, because they occur in the absence of a structure needing repair, since the relative clause without the pronoun is perfectly well-formed.

In fact, Prince presented examples in which the intrusive nominal is not a pronoun or even an epithet. She showed that any kind of nominal that can be anaphoric in discourse can serve the role of an intrusive pronoun (Prince 1990: (34a–d)):

(7) I had a handout and notes from her talk that that was lost too.

(8) He’s got this lifelong friend who he takes money from the parish to give to this lifelong friend.

(9) I have a manager, Joe Scandolo, who we’ve been together over twenty years.

(10) You assigned me to a paper which I don’t know anything about the subject.

These data support Prince’s contention that intrusive pronouns are not bound variable pronouns, unlike grammaticized resumptives, thus also supporting Sells’s original treatment of intrusive pronouns as discourse pronouns (Sells 1984).

Creswell (2002) rejects Kroch’s proposal that it is avoidance of ungrammaticality due to poor planning that leads to insertion of intrusive nominals. She contends that the grammar generates intrusive nominals in islands, based on theoretical considerations of Tree Adjoining Grammar (TAG; Joshi et al. 1975). Creswell observes that the theory of Frank (2002), which embeds TAG within the Minimalist Program (Chomsky 1995), does not permit generation of island-violating trees. She also notes that TAG-based models of incremental production (Ferreira 2000, Frank and Badecker 2001) do not permit Kroch’s solution, because they assume that the processing system uses only sub-trees licensed by the grammar to build larger trees and there are no such sub-trees in the case of island violations. Therefore, based on Creswell’s TAG assumptions, the underlying structure that the speaker would have to remedy, according to Kroch’s (1981) theory, is unavailable. Creswell (2002) argues that in fact the grammars of English speakers must independently have the resources required to form island structures containing intrusive
nominals. In other words, she argues that at least certain intrusive nominals are underlingly grammatical in English.

Ferreira and Swets tested the production of intrusive pronouns in \textit{wh}-islands by native speakers of English (Swets and Ferreira 2003, Ferreira and Swets 2005). They used a self-paced experimental task in which subjects were required to complete (in full sentences) partial descriptions that were presented with a picture array. The target sentences were sentences such as:

\begin{enumerate}
\item[(11)] This is a donkey that I don’t know where it lives.
\end{enumerate}

I will henceforth call this target sentence type “island-intrusive sentences” or “island-intrusives”.

Two control targets were also elicited. One controlled for surface length:

\begin{enumerate}
\item[(12)] This is a donkey that doesn’t know where it lives.
\end{enumerate}

The other controlled for length of the \textit{wh}-dependency without an island violation:

\begin{enumerate}
\item[(13)] This is a donkey that I didn’t say lives in Brazil.
\end{enumerate}

Ferreira and Swets conducted two versions of the production experiment. In the first experiment, subjects were allowed as much time as they wanted to plan the utterance before speaking. In the second experiment, a deadline procedure (Ferreira and Swets 2002) was used to pressure subjects to begin speaking quickly. If intrusive pronouns in \textit{wh}-islands are a result of lack of planning, as per Kroch (1981), then speakers in the no-deadline condition should plan the utterance to avoid both the island violation and the intrusive pronoun. For example, a participant could construct the following sentence instead of (11):

\begin{enumerate}
\item[(14)] This is a donkey and I don’t know where it lives.
\end{enumerate}

In contrast, in the deadline condition, speakers should not have time to plan an utterance that avoids an intrusive pronoun. Thus, the expectation is that, if intrusive pronouns are produced due to poor planning, there should be greater occurrence of intrusives in the deadline condition than in the no-deadline condition.

Instead, experimental participants overwhelmingly produced island violations like (11) in both conditions. In the no-deadline experiment, subjects could take as much time as they needed to plan their utterance before speaking and they did use this time, taking on average more than 2200 ms to begin. Nevertheless, 67\% of the sentences produced for the target condition involved an island containing an intrusive (Ferreira and Swets 2005: 274). In short, subjects did not use the extra time in the no-deadline experiment to plan an utterance that avoids an intrusive pronoun. In fact, the deadline experiment had a lower proportion, 56\%, of island-intrusive sentences. The deadline experiment did result in a greater number of alternative well-formed sentences that were not targets, like the coordination \textit{This is a donkey and I don’t know where it lives} or the left dislocation \textit{This donkey I don’t know where it lives}. Such alternatives were produced more often in the deadline experiment than in the no-deadline experiment, which leads Ferreira and Swets to conclude that island-intrusives are not mistakes. Such sentences are valid productions of the processing system (Ferreira and Swets 2005: 274).

Thus, one of Ferreira and Swets’s (2005: 274) main conclusions is that island-intrusives are “intentionally produced”; i.e. speakers plan to produce island-intrusive sentences. Swets and Ferreira (2003) outline a Tree Adjoining Grammar solution for generating the island-intrusive sentences and speculate that the reason that the structures are rejected despite being produced is that “the production and comprehension systems may set different parameters for accepting these structures.” The proposal is that the elementary trees required for producing the island-intrusives are part of the grammar and that the grammar treats island-intrusive sentences as well-formed for production, but not for comprehension. Like Creswell (2002), Swets and Ferreira (2003) contend that the island structures containing intrusive nominals are in fact underlyingly grammatical and they provide experimental evidence that they are not the result of poor planning. In their subsequent work, Ferreira and Swets (2005) no longer make the claim that the island structures are grammatical, and conclude only that “the production system has a modular architecture, because the message level system does not consider grammatical constraints when it creates a message level representation for the grammatical encoder” (Ferreira and Swets 2005: 276).

The claim that intrusive nominals in islands are grammatical is problematic. There are experimental studies that support the long-standing theoretical claim that the grammar of English does not license resumptive pronouns. Ferreira and Swets themselves ran two grammaticality judgment experiments. The first experiment was in the visual modality and the second in the auditory. Participants were asked to rate sentences like the island-intrusive (11) and the structural control (12) on a forced scale of 1 (perfect) to
5 (awful) based on whether the sentences were acceptable in English. In the visual presentation, island-intrusive sentences were rated at an average of 3.3, but control sentences rated at an average of 1.9. This significant difference was confirmed by the auditory presentation, where the average ratings were 3.0 (intrusives) and 1.7 (control), also a significant difference. These findings are all the more striking, since the participants, who rate the island-intrusive sentences as quite bad in comprehension, are from the very same population that reliably produced them.

These findings reflect other findings in the experimental literature. McDaniel and Cowart (1999) and Alexopoulou and Keller (2002, 2007) studied intrusive pronouns in islands using Magnitude Estimation (Stevens 1956, 1975, Bard et al. 1996). ME allows subjects to construct their own scale instead of using a forced scale. These studies found that an intrusive pronoun does not improve the grammaticality of islands. Alexopoulou and Keller (2002, 2007) demonstrate that, even at two levels of embedding, an intrusive does not improve the grammaticality of weak or even strong islands. These results have recently been replicated for complex noun phrase islands and relative clause islands by Xiang et al. (2008). Another recent study has shown once again that resumptives in English do not render a sentence more acceptable than a version with a gap, but that the acceptability of intrusive pronouns does improve in contexts that are harder to process (Hofmeister and Norcliffe 2013).

This brings us back to the ‘Intrusive Pronoun Puzzle’, which I will henceforth call the ‘Intrusive Nominal Puzzle’ in light of the observations, discussed above, by Kroch (1981) and Prince (1990) that the intrusive is not necessarily a pronoun:

(15) **Intrusive Nominal Puzzle (INP):**

Native speakers of English reject intrusive nominals as ungrammatical. Even in islands, intrusive nominals are as unacceptable as gaps.

Nevertheless, speakers produce intrusive nominals, even when they have time to plan alternative utterances without intrusives.

The proposal that there is a distinction between production and comprehension of intrusives is not appealing, because it posits distinct grammars for production and comprehension. This is not only unparasimonious, but also empirically and theoretically problematic, since it creates a much harder explanatory problem: that of explaining why the two systems are generally so congruent. Furthermore, other cases of underlying grammaticality, despite unacceptability, such as center embedding, do not similarly separate production and comprehension. In centre embedding, the unacceptable-but-grammatical form is also not produced, due to the same underlying performance restrictions that render it unacceptable in comprehension.

In what follows, I begin by reviewing the solution to the INP presented in Asudeh (2011a, 2012). The explanation takes facts at face value: intrusive nominal sentences in English are in fact underlyingly ungrammatical, but the production system prioritizes local well-formedness over global well-formedness in such a way that the sentences can nonetheless be produced. The solution is based on monotonic structure-building in LFG and has implications for directionality. The case of intrusives in islands in particular lends support to capturing islands from the outside in, since this allows the island structure to be produced, given the incrementality of production and the prioritization of local well-formedness. I first turn to an overview of some relevant concepts of LFG, before returning to the production model in section 4.

3. **Lexical-Functional Grammar**

3.1. **General Background**

Lexical-Functional Grammar is a lexicalist, constraint-based theory of grammar (Kaplan and Bresnan 1982, Bresnan 2001, Dalrymple 2001). Lexical items constitute key loci for grammatical information and for cross-linguistic variation. LFG posits two syntactic representations: c(ontext)-structure and f(unctional)-structure. C-structure represents word order, dominance and constituency, as modelled by a standard (non-tangled) tree; see the left side of (16). F-structure models more abstract aspects of syntax, such as predications and grammatical functions, null pronominals, local and unbounded dependencies, etc. F-structure is modelled as a feature structure; see the right side of (16). The φ correspondence function maps elements of c-structure to elements of f-structure, as exemplified in (16).
C-structure and f-structure have been understood for some time as just two components of a larger grammatical architecture, the Correspondence Architecture (Kaplan 1987, 1989, Halvorsen and Kaplan 1988, Asudeh 2004, 2006, Asudeh and Toivonen 2009, Asudeh 2012), which divides the form-meaning mapping into a series of simultaneously-present, discrete modules, each of which represents distinct linguistic information. One recent version of the architecture is shown in Figure 1. The various correspondence functions thus allow a single lexical entry to simultaneously specify constraints about a variety of grammatical information. The correspondence functions are functions in the mathematical sense and can therefore be composed (Kaplan 1987). For example, the original $\phi$ correspondence function between c-structure and f-structure is the composition of the $\lambda$ and $\alpha$ functions in Figure 1 (Butt et al. 1997).

$$\Gamma = \omega \circ \iota \circ \sigma \circ \lambda \circ \alpha \circ \rho \circ \mu \circ \pi$$

Figure 1: The Correspondence Architecture, pipeline version (Asudeh 2012)

LFG distinguishes the formal structures themselves from constraints on structures, which are structural descriptions that well-formed structures must satisfy. The formal structures of c-structure are phrase structure trees. Constraints on c-structures are phrase structure rules, where the right hand side of the rule is a regular expression (Kleene 1956). C-structure constraints in LFG thus support optionality, disjunction, conjunction, and unbounded repetition. The formal structures of f-structure are feature structures (a.k.a. attribute-value matrices). An f-structure is a finite set of attribute-value pairs. An attribute is a symbol and its value is either 1) a symbol, 2) a semantic form (a possibly complex symbol in single quotes), 3) a set, or 4) an f-structure. Constraints on f-structures are stated in a quantifier-free theory of equality. For example, the defining equation in (17) states that the f-structure $f$ has an attribute CASE whose value is ACCUSATIVE.

$$f\text{CASE} = \text{ACCUSATIVE}$$

Equation (17) is a defining equation, which defines the f-structure as having the attribute-value pair. An alternative sort of equation is a constraining equation (notated with a subscript $c$ on the equality: $=_{c}$), which checks that the minimal f-structure satisfying all defining equations has the attribute-value pair, but which does not itself result in the f-structure having the information. Other varieties of constraints are
negative equations, existential constraints and negative existential constraints. A comprehensive overview of LFG’s syntactic structures and constraints on these structures can be found in Dalrymple (2001: 7–176). For briefer overviews, see Dalrymple (2006) or Asudeh and Toivonen (2009).

3.2. Unbounded Dependencies

The original LFG theory of unbounded dependencies depended on special meta-variables in c-structure (Kaplan and Bresnan 1982: 231–263). The phrase-structural approach was abandoned in light of evidence and argument that unbounded dependencies are subject to functional constraints and should therefore be captured at f-structure (Kaplan and Zaenen 1989). The formal mechanism that allows this is functional uncertainty over grammatical functions in constraints on f-structures. The constraint language is a regular language, so the uncertainty is formally captured through regular operations, such as disjunction, negation, and the Kleene operators.

F-structure constraints for filler-gap dependencies in the Kaplan and Zaenen (1989) theory have the following general form:

\[(\uparrow \text{Top}) = (\uparrow \text{Body Base})\]

The top of the dependency, schematized as \textbf{Top}, is normally taken to be one of the discourse functions \textit{TOPIC} or \textit{FOCUS} (Kaplan and Zaenen 1989, King 1995, Bresnan 2001), as in f-structure (16) above. There is some theoretical motivation for eliminating \textit{TOPIC} and \textit{FOCUS} as syntactic entities (Asudeh 2011b, 2012), but I will not argue for that here. I will henceforth assume a single \textit{UNBOUNDED DEPENDENCY FUNCTION} (UDF), just as an expedient simplification.

A simple functional uncertainty is the disjunction in (19), which states that the \textbf{Base} of the unbounded dependency may be a \textbf{SUBJ} or \textbf{OBJ}.

\[ (\uparrow \text{UDF}) = (\uparrow \text{Body} \ [\text{SUBJ} \mid \text{OBJ}]) \]

This equation allows the unbounded dependency to terminate in either a \textit{SUBJECT} or an \textit{OBJECT} by using disjunction to model uncertainty about the functional identity of \textbf{Base}. This kind of disjunction is routinely generalized by defining an f-structure metacategory, GF (‘grammatical function’), as the disjunction of all grammatical functions in the grammar (Dalrymple 2001: 140). This sort of uncertainty can be further fine-tuned using the regular language complementation operator, \(\sim\). For example, if a grammar allows extraction of all grammatical functions except obliques and adjuncts, then \textbf{Base} could be written as \([\text{GF} \sim \{\text{OBL} \mid \text{ADJ}\}]\).

Functional uncertainty of the form in (19) does not express unboundedness. This is expressed by using the Kleene star and plus operators, where Kleene * has the standard interpretation “zero or more” and Kleene + has the standard interpretation “one or more”. Here is an example:

\[ (\uparrow \text{UDF}) = (\uparrow [\text{COMP} \mid \text{XCOMP}]^* \text{ Base}) \]

This equation allows the unbounded dependency to pass through any sequence of COMP and XCOMP f-structures (including none).

A noteworthy consequence of this treatment of unbounded dependencies is that no empty category needs to be posited in the syntax. The extracted constituent’s base position is absent entirely in c-structure. At f-structure, the UDF grammatical function borne by the \textbf{Top} of the dependency is equated with the grammatical function of the \textbf{Base}. This is illustrated in example (16) above: there is no c-structural representation of the missing embedded subject and the relevant \textbf{SUBJ} grammatical function in f-structure is equated with the unbounded dependency function \textit{FOCUS} contributed by the \textit{wh}-phrase. The Kaplan and Zaenen (1989) theory thus removes extraction per se as an argument for empty categories. However, although empty categories are not necessary in c-structure, the theory does not formally preclude them, should there be some independent motivation (e.g., if Bresnan’s arguments from weak crossover are accepted (Bresnan 1995); but see Dalrymple et al. 2001, 2007, Nadathur 2013).

The Kaplan and Zaenen unbounded dependency constraints are based on a generalization of standard LFG functional application, which is outside-in, or concerns the characterization of embedded f-structures. For example, the vanilla constraint \((\uparrow \text{TENSE}) = \text{PAST}\) states that the f-structure designated by \(\uparrow\), which is a function, applied to the argument \text{TENSE} has value \text{PAST}; i.e. \((\text{TENSE}, \text{PAST}) \in \uparrow\). This means that the attribute-value pair in question is embedded in f-structure \(\uparrow\). Outside-in functional application can be iterated to characterize paths through embedded f-structures. For example, the standard LFG functional control equation (Bresnan 1982), \((\uparrow \text{SUBJ}) = (\uparrow \text{XCOMP SUBJ})\), equates the \text{SUBJ} of \(\uparrow\) with a
SUBJ embedded inside an XCOMP that is in turn embedded in ↑. Lastly, iterated outside-in functional application is generalized to outside-in functional uncertainty by characterizing paths from outer f-structures to inner f-structures as a succession of outside-in functional applications, where uncertainty over paths can be expressed using the regular language of f-structure constraints.

Bresnan (1995, 2001) presents an alternative theory of unbounded dependencies in which there is an empty category at the Base of the dependency. The empty category is associated with an inside-out equation, a different sort of f-structure constraint. The general format of inside-out unbounded dependency constraints is:

(21) \(((\text{Body Base} ↑) \text{Top}) = ↑\)

The equivalent constraint to the particular constraint (20) on the Bresnan theory would be written as follows:

(22) \(((\text{COMP} | \text{XCOMP})^* \text{ Base} ↑) \text{UDF}) = ↑\)

Constraint (22) associates the f-structure of the empty category that bears the constraint with the UDF of an f-structure that contains the empty category’s f-structure. The outer f-structure must be reachable from the inner f-structure that corresponds to the empty category by traversing only COMP and XCOMP f-structures.

Inside-out expressions like (22) were first introduced in unpublished work by Kaplan and first discussed in print by Halvorsen and Kaplan (1988). Inside-out expressions concern the characterization of f-structures that contain the f-structure that is the locus of the constraint. For example, the constraint (SUBJ ↑) states that ↑ must be the value of the attribute SUBJ in the f-structure that contains ↑; i.e., if ↑ is f and the f-structure that contains f is g, then \{SUBJ, f\} ∈ g. Inside-out expressions can be iterated to characterize paths through enclosing f-structures, as in (22). Iterated inside-out expressions can be generalized to inside-out functional uncertainty by characterizing paths from inner f-structures to outer f-structures as a succession of inside-out computations, where uncertainty over paths can be expressed using the regular language of f-structure constraints. In addition to its role in Bresnan’s treatment of unbounded dependencies, inside-out expressions are the key formal mechanism in Dalrymple’s theory of the syntax of anaphoric dependencies (Dalrymple 1993) and Nordlinger’s theory of case (Nordlinger 1998).

Island constraints (Ross 1967) are captured through constraints on Body and Base. For example, the constraints in (23) (outside-in) and (24) (inside-out) equally capture both the Sentential Subject Constraint and the Left Branch Condition (Ross 1967).

(23) \((↑ \text{UDF}) = (↑ \{\text{XCOMP} \mid \text{COMP}\}^* \{\text{SUBJ} \mid \text{OBJ}\})\)

(24) \(((\{\text{XCOMP} \mid \text{COMP}\}^* \{\text{SUBJ} \mid \text{OBJ}\} ↑) \text{UDF}) = ↑\)

The Sentential Subject Constraint is captured because Body does not contain SUBJ; the unbounded dependency cannot pass through a SUBJ f-structure, only through an XCOMP or COMP. The Left Branch Condition is captured, because the Base functions SUBJ and OBJ cannot be functions borne by possessors, for independent reasons.

Constraints on extraction are also captured through off-path constraints (Dalrymple 1993), which place restrictions on f-structures found along the path through Body and into Base. Off-path constraints can be defined as in (21), following Dalrymple (2001: 151):

(25) a. In an expression \(\langle\langle \leftarrow a \rangle \rangle\), \langle\langle \leftarrow\rangle \rangle refers to the f-structure of which a is an attribute.

b. In an expression \(\langle\langle \rightarrow a \rangle \rangle\), \langle\langle \rightarrow\rangle \rangle refers to the value of the attribute a.

For example, wh-islands can be captured with an equation like the following, which states that no GF on the path to the base can have its own UDF function:

(26) \((↑ \text{UDF}) = (↑ \text{GF}^* \text{UDF} \rightarrow \langle\langle \rightarrow\rangle \rangle)\)

The wh-island violation example (27) has the f-structure in (28).

(27) *Who did John wonder what Mary gave?

2 This discussion is based on Asudeh (2011b), where the issues are discussed in more detail.
The first GF that the unbounded dependency passes through is the matrix COMP. This COMP itself contains a UDF, due to the lower extraction of what. The structure is therefore excluded if the grammar contains constraint (26).

4. Outside-In Uncertainty and Prioritization of Local Well-Formedness

I have proposed the simplified production model in Figure 2 (Asudeh 2011a, 2012), based on the assumption that sentence production is incremental (Kempen and Hoenkamp 1987, Levelt 1989). The incrementality of the model is based on LFG’s model of the “fragmentability of language” (Bresnan 2001: 79–81), which is related to the monotonicity property.

Bresnan notes that LFG can distinguish between informative and uninformative sentence fragments, using the f-structure representation of predicate-argument relations. The fragment . . . seems to . . . , as in (29), contrasts with the fragment . . . to by for . . . , as in (30).

(29) [Speaker A:] And he agrees?
    [Speaker B:] — seems to.

(30) The one he should be spoken to by, for God’s sake, is his mother.

The first fragment constructs an informative partial c-structure and f-structure, which form subparts of the c-structure and f-structure for a full sentence like He seems to agree. The second fragment constructs only three unrelated structures (Bresnan 2001: 81). The ability to define informative fragments arises from the fact that the c-structure head/f-structure predicator contains information about the larger structures in which it can be embedded. In particular, the f-description (set of constraints) of the lexical items in the fragment defines its f-structure. In the case of fragment (29), the lexical items for seems and to in (31) construct the partial c-structure and f-structure in (32) (based on Bresnan 2001: 80–81).
seems

(↑ PRED) = 'seem⟨XCOMP⟩SUBJ'

(↑ SUBJ) = (↑ XCOMP SUBJ)

(↑ TENSE) = PRES

(↑ SUBJ PERSON) = 3

(↑ SUBJ NUMBER) = SING


PRED

'seem⟨XCOMP⟩SUBJ'

SUBJ

PERSON 3

NUMBER SING

XCOMP

SUBJ

TENSE

PRES

The monotonicity property dictates that the f-structure of the fragment must subsume, or be more general than, the f-structure of any larger sentence in which it might have been embedded. In other words, monotonicity restricts which sentences/propositions may be recovered from a fragment.

Levelt (1989: 258) argues that the planning units of sentence production are not determined by syntactic criteria, but rather reflect a “function-argument structure”, which he takes to be semantic in nature. The function-argument structure is something like a rough thematic structure, akin to the one offered by the Conceptual Semantics of Jackendoff (1990, 1997, among others). In order to make this more precise, I assume that when a speaker begins initial planning s/he puts together a message that identifies the event or state, its basic function/predicate, the function’s arguments and their rough thematic relation to each other. The speaker then identifies what sort of utterance s/he wants to make with respect to these elements: a declaration, a question, etc. The thematic structure of the message unfolds through incremental construction of fragments of grammatical structure. These are monotically added to the grammatical structure computed thus far (in the absence of explicit repairs). Incremental grammatical production is ultimately based on the predicate-argument structure of heads, which is lexically encoded and will bear a close relationship to the function/argument structure of the planning unit.

LFG is a declarative, constraint-based theory and therefore does not have a notion of procedural grammatical generation. However, incremental processing is inherently procedural and involves issues of timing and ordering of operations. Procedurality of production and parsing and timing of grammatical operations are central questions in psycholinguistics (for an overview, see Frazier 1999). The model presented here captures procedurality of production through chunking. Each successive chunk of grammatical representation produced by the processing system must be grammatical in its own right — locally well-formed — in order to be generated. This leads to incremental generation of a grammatical structure that satisfies local grammaticality requirements at each step. However, the end result does not necessarily satisfy global grammaticality. Global well-formedness depends on successful monotonic integration of subsequent fragments, which may fail even though the fragments computed thus far are each locally well-formed.

If we assume the outside-in theory of unbounded dependencies, given the incrementality of production, there is an important consequence for the construction of locally well-formed grammatical representations. When the chunk that contains the unbounded dependency is under construction for production, the top of the unbounded dependency contributes the outside-in equation. This equation is satisfied only if an otherwise empty grammatical function is found in a subsequent, embedded f-structured. If the unbounded dependency grammatical function and an embedded grammatical function are equated, the filler at the top of the dependency and what we might call the gap at the base are successfully integrated. However, the term ‘gap’ in this context is a purely pre-theoretical one: there is no element in c-structure or f-structure that corresponds to the gap. Just as what we call a gap is in the sentence is identified by the absence of local material, the only structural correlate of the absent local material is absence of structure. In other words, the gap is just nothing.

This has an important implication for incremental construction of fragments. The outside-in function contributed by the filler is unbounded and defines a path through f-structure material that is still being incrementally constructed. If the grammar cannot integrate the filler into the local f-structure being constructed because all grammatical functions are locally filled, it does not crash, because the integration site
could be in the next chunk of f-structure that is yet to be constructed or in a chunk after that. The system has the following properties:

1. The functional uncertainty equation is unbounded.
2. The equation is initiated at the top of the unbounded dependency, which corresponds to an earlier chunk in production.
3. The gap is not marked in the locally produced c-structure or f-structure.

Given these properties, it is possible to maintain a model of incremental processing in which integration of the filler by the grammar takes place after the local structure under construction has been built.

In constructing a local structure, each grammatical function (GF) can be handled by the production system in one of two ways. First, the production system can leave the GF empty; the GF must then be subsequently licensed by integration of a filler. The filler must be functionally equated with this GF before production of the next chunk, or else the local structure of the current chunk would not be well-formed. Alternatively, the production system could posit lexical material, such as a pronoun or other nominal, that will fill the GF and that is consistent with the other specifications of the local structure. For example, in English, an OBJECT pronoun must have accusative case. The lexical material that is chosen to fulfill the local requirements for the GF must also be consistent with the current message plan. If this second option is chosen, the filler is not integrated but the local structure is well-formed. The filler must instead be integrated in an upcoming chunk.

The two situations are sketched in Figures 3 and 4. Figure 3 shows what happens when the filler is integrated. Figure 4 shows what happens when the filler is passed through the local structure rather than being integrated. The second pattern is the relevant one for explaining how ungrammatical intrusives are produced instead of a gap.

5. A Processing-Based Explanation of English Intrusive Nominals

I now turn to two examples to better illustrate how the model of Asudeh (2011a, 2012) solves the Intrusive Nominal Puzzle, both in island and non-island contexts. I then return to the question of outside-in versus inside-out functional uncertainty in section 6.

5.1. Example 1: An Intrusive Where a Gap Would be Well-Formed

The first example is the Prince (1990) example in (6), which is syntactically quite simple. This not only makes it suitable for illustrative purposes, it more importantly demonstrates that this account of production does not appeal to processing complexity to explain the Prince and Kroch examples. Example (6) is repeated here:

(33) You get a rack that the bike will sit on it.

Initiation of production is shown in (34). The local structure under construction is indicated by the dashed box. I use a common LFG convention of abbreviating the contents of f-structures with the contributing lexical item in double quotes; e.g. “you” represents PRED ‘pro’ and PERSON 2. The planned message is represented in the ellipse. It is not normally the case that an entire utterance is planned in advance (that would not be a Levelt-style model). However, the findings of Ferreira and Swets (2005) indicate that the production system plans at least far enough in advance to include a message of this length and complexity in the initial plan.

3 The interaction of complexity and intrusive pronouns is discussed in Asudeh (2012).
Figure 3: Local well-formedness satisfied by integration of a filler

Figure 4: Local well-formedness satisfied by insertion of lexical material
The first fragment is made up of the head *get* and its arguments. I assume that relative clause construction begins at this stage, too. This seems reasonable since the relativizer can be prosodically grouped with the relative head (in contrast to a non-restrictive relative). This chunk is locally well-formed: all of *get*'s arguments are present.

The relative clause launches an unbounded dependency. This is represented by the UDF in the innermost f-structure *f* in (34), where *f* is an arbitrary label for the relevant f-structure. The UDF is associated with an outside-in constraint like the following:

\[(↑ UDF) = (↑ . . . GF)\]

The up arrow meta-variable in the outside-in functional uncertainty is set to *f*. The ellipsis indicates the Body of the unbounded dependency, which is material yet to be generated. The Base of the unbounded dependency is represented simply as GF for the purpose of this illustration. The functional uncertainty that is initiated by the relative clause is carried over to construction of the next chunk.

In constructing the next chunk, the production system can go for either of the options outlined in Figure 3 and Figure 4. If the first option is chosen, the filler is integrated into the local structure being constructed; the relative clause is constructed with a gap:

\[(36) \quad \text{You get a rack that the bike will sit on.}\]

The construction of the local structure is shown here:
The filler is integrated into the local structure, satisfying both the demands of the filler and the local demand that the OBJ must be integrated into the f-structure. The overall construction of the sentence is illustrated here:

(38) You get a rack that . . . 
✓ Locally well-formed

✓ Locally well-formed

✓ Globally well-formed

Each of the local structures here is well-formed and consistent with the plan and the overall result is also well-formed.

Chunk 2 can alternatively be constructed through the option of inserting lexical material that is consistent with the plan, as sketched in Figure 4. In this case, the Prince example (33) is produced instead. The local structure under construction is again shown in the dashed box. The pronoun it has been inserted as the prepositional object.
The production system passes the filler on and attempts to continue, but there is no longer anywhere to integrate the filler. There is no remaining structure to be built and insertion of the filler in the structure built so far is impossible. The situation is illustrated in (40):

(40) You get a rack that . . . Filler … the bike will sit on it Filler

✓ Locally well-formed

* Globally ill-formed

The grammar ultimately fails to license the structure that has been attempted. However, due to incremental production of locally well-formed structures, the ungrammatical sentence has been produced; i.e. it is uttered. Incremental production nevertheless treats each stage of producing (33) as locally grammatical. The result of production is, however, globally ungrammatical and is perceived as such by native speakers. The perception of ungrammaticality does not arise through production, but rather through parsing. What the parser does with the result of productions like (33) is considered in detail in Asudeh (2012).

The question arises of what constraints there are on the construction of locally well-formed grammatical structures through lexical filling (Figure 4) — which is what leads to the construction of sentences like You get a rack that the bike will sit on it. If there are no grammatical constraints, then this sentence and other intrusive nominal sentences are effectively speech errors. I do not think the relevant sentences are speech errors in this sense. First, they are constrained at the level of local grammatical structure by the kinds of local structure that can be well-formed. For example, in constructing the sentence You get a rack that the bike will sit on it, insertion of a pronoun as the object of on is locally licensed by the rule that constructs PPs, the lexical requirements of on which require an OBJ, the fact that the OBJ of on must be realized by a nominal, etc. If local grammatical well-formedness is a criterion, then speakers could not instead produce things like You get a rack that the bike will sit it. If this kind of form is produced at all, it really is a speech error. This must be distinguished from locally well-formed structures that arise from purely incremental production. Second, the kinds of things that can be inserted are constrained by the plan itself. If the speaker wants to say something about a rack, then s/he will select a lexical item that is consistent with that plan. In English, the kinds of lexical items that are consistent with the plan are pronouns (it), deictics (that), names and definite descriptions that refer to the thing in question (the rack), and epithets (the damn thing). Third, the findings of Ferreira and Swets (2005) indicate that speakers produce these utterances even when given time to plan carefully (no-deadline experiment). The utterances are thus more deliberate than speech errors.
5.2. Example 2: An Intrusive Where a Gap Would be Ill-Formed

The more interesting case is that of island violations like the Ferreira and Swets donkey example, repeated here, or Creswell’s attested example in (42).

(41) This is a donkey that I don’t know where it lives. (Swets and Ferreira 2003)

(42) You have the top 20% that are just doing incredible service, and then you have the group in the middle that a high percentage of those are giving you a good day’s work … (Creswell 2002: 102, (4d); http://www.ssa.gov/history/WEIKEL.html4)

The explanation of these cases basically reduces to the case that we have already looked at plus the fact that the island prevents integration of the filler.

Let us use the simpler donkey example as an illustration of island-intrusive cases. Production starts as follows:

(43)

TOPIC: donkey-X
DEIXIS: this-Y
BACKGROUND: speaker-S, hearer-H
MESSAGE: Y is X. S does not know where X lives.

Chunk 1: This is a donkey that …

Lexicon

\[
\begin{align*}
\text{PRED} & \quad \text{‘be’} \\
\text{SUBJ} & \quad \text{“this”} \\
\text{OBJ} & \quad \text{“a”} \\
\text{ADJ} & \quad \{ f[UDF \text{ “pro”}] \}
\end{align*}
\]

\( \checkmark \) Locally well-formed

An unbounded dependency is once again launched by the relativizer. Let us assume that the \( \text{wh} \)-island constraint is stated as an off-path equation to the effect that the functional uncertainty cannot pass through a \( \text{COMP} \) that contains a \( \text{UDF} \), as discussed in section 3.2. A simplified version of the functional uncertainty is shown here:

(44) \( (↑ \text{UDF}) = (↑ \text{COMP}^* \text{GF}) \rightarrow (\rightarrow \text{UDF}) \)

The equation states that the grammatical function to be equated with the \( \text{UDF} \) can be found by going through zero or more \( \text{COMP} \) f-structures and none of the \( \text{COMP} \) f-structures may have an unbounded dependency function (UDF) of its own. After construction of the first chunk, the \( \text{UDF} \) has not been integrated and the beginning of the path has already been instantiated to one \( \text{COMP} \).

In producing the next chunk,\(^5\) the production system constructs the following partial local structure (indicated by the dotted box):


\(^5\) I assume for simplicity that the next chunk is the remainder of the sentence. Nothing depends on this.
(45) \[\text{Chunk 2: \ldots I don’t know where it lives}\]

\[
\begin{aligned}
\text{Filler:} & \quad (f \quad \text{COMP} \quad \ldots) \\
\text{\neg} & \quad (\rightarrow \text{UDF})
\end{aligned}
\]

In this chunk, the production system still has an unintegrated unbounded dependency and now it has encountered a new one. The option of positing a gap for the most deeply embedded SUBJ, as in Figure 3, does not exist. The presence of the embedded UDF means that there is no way to locally satisfy the relative clause’s functional uncertainty equation. There is also no global way to satisfy the equation. As soon as a COMP containing a UDF is encountered, satisfaction is impossible. The result is that the only way to construct a locally well-formed f-structure is through the option in Figure 4 of inserting some lexical material that is consistent with the plan (i.e., that refers to the donkey). The filler does not pass through the chunk, though, because it cannot do so and satisfy the unbounded dependency equation. The new unbounded dependency also needs to be integrated, but this can be done using the gap integration option in Figure 3, because there is no embedded in UDF in the way of the lower integration.

The final local structure is shown here in the dashed box:
The grammatical architecture of LFG and the production model proposed in Asudeh (2011a, 2012), reviewed here, offers a solution to the Intrusive Nominal Puzzle. Sentences containing intrusive nominals are deemed ungrammatical by native speakers, because they are underlyingly ungrammatical. Nevertheless, the sentences in question can be produced by these speakers through the interaction of fundamental properties of the overall system. Incremental production is modelled as monotonic growth of syntactic structure through integration of fragments. Each fragment must be locally well-formed. Local well-formedness depends in part on successful insertion of all grammatical functions. The two strategies for the latter are insertion of lexical material or equating the grammatical function with a filler that has previously been introduced. Insertion of lexical material can guarantee local well-formedness in cases where filler integration is impossible (islands).

5.3. Summary

The explanation of intrusive nominal productions rests partly on a prioritization of local well-formedness over global well-formedness. Another important aspect of the explanation is that the unbounded dependency’s functional uncertainty equation is outside-in. If the production system is working with an outside-in unbounded dependency equation, then it is not determinate that any given local structure is the correct integration site, because the site may be in an upcoming fragment yet to be constructed. Local
well-formedness is then achieved through the strategy of positing lexical material (rather than by integrating the top of the unbounded dependency), as sketched in Figure 4. In the case of an island, the local structure can only be locally well-formed using the lexical insertion strategy. Importantly, as the island is being constructed by the production system, it is already determinate that the unbounded dependency integration strategy is unavailable, due to the outside-in equation that has already been generated when the top of the dependency was produced.

Both the island and non-island situations for outside-in functional uncertainty contrast with what would have to logically be the case if the unbounded dependency were instead licensed by an inside-out functional uncertainty associated with a gap. In the non-island case, if a gap is posited, then the system will need to search all of the structure grown so far for an associated filler. This contrasts with the outside-in alternative, where the constraint can be checked successively as each chunk of structure is added. The final check when the gap is encountered is then effectively just a local check. The island case offers an additional point of divergence. In the inside-out alternative, the unbounded dependency constraint is associated with an empty category. The unbounded dependency constraint specifies all the licit ways in which the gap can be associated with a filler; by definition, island violations are illicit associations of filler and gap. Therefore, the fact that the local structure is an island is only registered by the grammar if the gap is posited. Therefore, on the inside-out alternative, it is not determinate that a gap would lead to local ill-formedness until the empty category corresponding to the gap is in fact generated. In short, this would mean, as far as the production system is concerned, that local well-formedness is equally plausible given either the filler-gap strategy (Figure 3) or the lexical insertion strategy (Figure 4). We would then expect to see roughly equal distributions of the two strategies in production of island sentences. This is emphatically not the case. For example, Swets and Ferreira (2003) report that the proportion of actual wh-island gaps produced in the wh-island condition (i.e., island-intrusive condition) were 1.4% for the no-deadline experiment and 3.1% for the deadline experiment.

The thrust of the outside-in/inside-out contrast remains even if island constraints are partly or wholly extra-grammatical in nature, as argued by Deane (1991), Kluender (1991, 1998), Kluender and Kutas (1993), and Hofmeister and Sag (2010), among others. If island constraints are not part of the grammar, then the unbounded dependency constraints would be simplified. Specifically, the off-path constraints would not be required and the Base of the dependency could possibly be unrestricted (i.e., fully general as to possible grammatical function of the gap). The important distinction between the two kinds of constraints is that the outside-in constraint can already be checked outside the island (however it is defined), whereas the inside-out constraint can only be checked once the gap has been posited in the island.

7. Conclusion

I have reviewed a model of the language faculty in which production is monotonically incremental (Asudeh 2011a, 2012). Crucially, with respect to directionality, I have argued that unbounded dependency constraints are generated outside-in (working in the direction less embedded to more embedded; i.e. from an outer f-structure to an inner f-structure), rather than inside-out (working in the direction more embedded to less embedded; i.e. from an inner f-structure to an outer f-structure). I also assumed that the processing system prioritizes local well-formedness over global well-formedness. Effectively, processing resources are concentrated on a relatively small piece of structure at a time.

This system allows a natural, although surprising, explanation of the Intrusive Nominal Puzzle. Intrusive nominals are produced because the structures that contain them are locally well-formed. However, speakers nevertheless judge them as ungrammatical, because English does not have the grammatical resources to properly license resumptive pronouns. As a result, sentences containing intrusive pronouns and other intrusive nominals are globally ill-formed. In other words, through the normal constraints of the production system, speakers are, in certain circumstances, monotonically and incrementally producing sentences that are underlying ungrammatical.

With respect to the issue of directionality, this paper makes two key points. The first is that there is a distinction in predicted outcomes for production of intrusive nominals in islands, depending on whether the unbounded dependency constraint can be checked outside of the island or only once inside the island. I argued that the outside-in alternative comports better with Ferreira and Swets’s production study. I do not think this is necessarily an LFG-specific argument, but rather a distinction between theories in which islands are identified internally to the island (e.g., Principles & Parameters Theory, TAG; Chomsky

\[\text{See Sprouse and Hornstein (2013) for a recent collection of psycholinguistically-motivated papers on this debate.}\]
1973, 1986, 2001, Frank 2002), versus those in which islands are or can be identified externally to the island (e.g., Categorial Grammar and HPSG; Steedman 1987, Morrill 1994, 2011, Pollard and Sag 1994, Bouma et al. 2001, Levine and Hukari 2006). In fact, the distinction seems relevant even if islands are completely or in part extra-grammatical, since the model relies on a mix of grammatical and processing factors. The second, more general point is that LFG’s outside-in and inside-out constraints are rich loci for investigation of questions of grammatical directionality, even in a framework that is non-derivational.

References


Do the “right” move
Cristiano Chesi
NeTS - IUSS Center for Neurolinguistics and Theoretical Syntax, Pavia
cristiano.chesi@iusspavia.it

In this paper I argue that rightward dependencies such as Extraposition and Heavy NP-Shift can be accommodated, cross-linguistically, in a top-down (Chesi 2004-12) minimalist framework that results in a left-right phrase structure generation (cf. Phillips 1996, Phillips and Lewis, Kempson et al. and Kiaer in this volume). All the special properties that make Extraposition and Heavy NP-Shift peculiar compared to standard “leftward” movement will be considered, focussing on clause-boundedness, adjunct/argument asymmetries with respect to dependency directionality, and the definiteness constraint.

1. Introduction
A Minimalist Grammar is top-down (Chesi 2004-12) and, as a consequence, left-right (cf. Phillips 1996, and contributions in this volume by Phillips and Lewis, Kempson et al. and Kiaer) if structure building procedures operate by expanding constituents from the root of the syntactic tree, rather than creating them by merging lexical items starting from inner constituents (Chomsky 1995). Reformulating the anti-symmetric intuition (Kayne 1994), I suggest that linear ordering (“left-right”, in the sense of the temporal order in which roughly we parse and produce sentences) derives from the hierarchical one (top-down). Surprisingly, at least from the mainstream perspective, in such a framework, all long-distance dependencies are “rightward” dependencies given that the dependency trigger must be found first (e.g. a wh- filler, Fodor 1978), then the dependent (possibly non-local) constituent will be identified (e.g. the selecting verbal head). Using memory buffers to store and retrieve constituents in a principled order/way we can correctly characterize (successive cyclic) movement and islandhood (Chesi 2004-07), parasitic gap constructions (Bianchi & Chesi 2008), quantifier raising (Bianchi & Chesi 2010) and A-binding (Bianchi 2010). Here I argue that in any (standard) theory that includes the notion of feature-driven movement, rightward movement (like PP and (restrictive) Relative Clause (rRC) Extraposition (EXT) in SVO languages) shows peculiar properties that are usually resistant to a non-stipulative unified account. These properties are “clause”-boundedness (1) ((1a) from Baltin 2006, (1b-b') from Akmajian 1975), adjunct/argument asymmetries sensitive to displacement directionality (2) and the definiteness constraint (3) (from Fox and Nissenbaum 1999):

(1) a. *[John was believed [ to be certain _ ] by everybody] that the Mets would lose.
   b. *[A review of [a book _ ] ] appeared by three authors
   b'. [A review [of a book] _ ] appeared by three authors

© Copyright 2013 Cristiano Chesi
STIL – Studies in Linguistics. 6:107-138
Special Issue on “Directionality of Phrase Structure Building” ISSN: 2281-3128
Do the “right” move

Cristiano Chesi

(2) a. We saw [a painting _] yesterday of John
   a'. Of whom did you see [a painting _]?
b. We saw [a painting _] yesterday from the museum
   b'. "? From where did you see [a painting _]?

(3) a. I saw the (best) picture yesterday from the museum.
   a'. ?I saw the (best) picture yesterday of the museum.
   a''. I saw a (very good) picture yesterday of the museum.

Here I want to show that this cluster of properties can be handled in a natural way if we drift away from the standard conception of “bottom-to-top” (in the sense of Phillips 1996) derivation and we redefine structure-building operations top-down within a phase-based
1, head-driven
2 derivation.

In the first part of this paper (§2), I will compare rightward movement in English, Dutch and Italian, summarizing some fundamental findings on Extraposition: what moves (§2.1), from where (§2.2) it does and where the extraposed element seems to be attached (§2.3).

In the second part (§3), I will discuss data related to Heavy NP-Shift that present properties consistent with Extraposition (i.e. Clause-Boundedness).

In the third part (§4), I will analyse some of the proposed standard solutions, highlighting the main problems of the given analyses.

The last section (§5) is dedicated to the proposed solution: the memory-based approach to long-distance dependencies (Movement, QR and Binding) seems to be able to accommodate most of the asymmetries and problems reported in these pages, and it provides a principled account of Extraposition and Heavy NP-Shift constraints without losing the ability to discriminate between classic “rightward” and “leftward” movement.

2. A case of rightward movement: Extraposition (EXT)

Extraposition (henceforth EXT) is a limited option that some language makes available to marginalize, at the end of the relevant clause, certain constituents (mainly PPs and rRCs). The well know distributional properties that characterize this phenomenon are summarized in this chapter, focussing on what moves (§2.1), from where it does (§2.2) and where the extraposed constituent attaches (§2.3).

2.1. What

As mentioned before ((2), repeated in (4)) we can observe that Movement and EXT show an asymmetric sensitivity to the argument/adjunct distinction: while EXT is insensitive to such opposition, sub-extraction from NPs, in English ((4), Fox and Nissenbaum 1999), Italian (5) and Dutch (6), show a pretty neat degradation if the extracted element is not a complement but an adjunct.

(4) a. We saw [a painting _] yesterday of John
    a'. Of whom did you see [a painting _]?

1 As in Chomsky 2008, here a phase will be a minimal computational domain bounding structure building operations, though this notion will be revisited according to the perspective shift.
2 As in HPSG (Pollard and Sag 1994) I believe that the lexical verb within CP and the main noun in any DP/PP plays a crucial role that must be acknowledged.
b. We saw [a painting _ ] yesterday from the museum
b’. *??From where did you see [a painting _ ]?

(5) a. pro Abbiamo visto [un quadro _ ] ieri di Gianni
    pro have seen a painting yesterday by G.
a’. Di chi pro hai visto [un quadro _ ] ieri?
   Of whom pro have seen a painting yesterday
b. pro Abbiamo visto [un quadro _ ] ieri con dei girasoli
    pro have seen a painting yesterday with sunflowers
b’. *Preso da dove / Da dove pro abbiamo visto [un quadro _ ]?
    taken from where / from where pro have seen a painting

(6) a. We hebben gisteren [een schilderij _ ] gezien van Jan.
    We have yesterday a painting seen of J.
a’. Van wie heb jij [een schilderij _ ] gezien?
    Of whom have you a painting seen
b. We hebben gisteren [sinaasappels _ ] gekocht uit Spanje
    We have yesterday oranges bought from Spain
b’. *Waaruit / waar vandaan heb jij gisteren [sinaasappels _ ] gekocht?
    Where / from where have you yesterday oranges bought?

Also restrictive relative clauses undergo Extraposition (English (7), De Vries 2006, Italian (8), and Dutch (9), De Vries 2002)3:

(7) [Some men _ ] appeared at the door that Mary had been insulting.
(8) Ho visto [un uomo _ ] oggi che aveva perso la sua valigia.
    (I) have seen [a man _ ] today who have lost the his bag
(9) Ik heb [de man _ ] gezien die zijn tas verloor.
    I have [the man _ ] seen who his bag lost.

Briefly:
   i.   EXT, against movement, is insensitive to the argument/adjunct distinction;
   ii.  PPs and rRCs can be extraposed and no crucial cross-linguistic variation
        (among English, Italian and Dutch) seems to hinge on that.

2.2. From where
Following Baltin (2006) let us call the position “related” to the extraposed constituent
(i.e. the position “from where” EXT takes place) the host position. A host can be both
an argument (10a,b,c) and an adjunct (10d):

(10)a. [A book _ ] appeared which was written by Chomsky
    b. I called [somebody _ ] yesterday who I couldn't stand
    c. I talked [to somebody _ ] about that who was quite knowledgeable
    d. I saw it [in a magazine _ ] yesterday which was lying on the table

3 The extraposition of appositive relative clauses is a more controversial issue in the linguistic
this paper, I will restrict the discussion to restrictive relative clauses only.
Here some restrictions apply if the constituent is moved: Fronted PPs adjuncts (11a) do not permit EXT to take place (even though prepositional stranding improves EXT from these constituents, (11a’)), while moved wh- arguments (e.g. (11b), Baltin 2006) do.

(11) a. *[In which magazine _] did you see it \( t_i \) which was lying on the table?
a’. **[Which magazine _] did you see it in \( t_i \) which was lying on the table? 
b. [Who _] did you visit \( t_i \) who was unhappy about the visit?

While in languages like English or Dutch, Extraposition from any argument/adjunct position is generally accepted, this is not the case for languages like Italian, where it is not possible to extrapolate a PP/rRC from the pre-verbal subject position (independently of whether the verb is transitive (12a), unergative (12b), unaccusative (12c), or in a passive voice (12d)):

(12) a. *[Un amico _] ha raccontato questa storia di Gianni / che ho visto ieri
   A friend has told this story of G. / which (I) have seen yesterday
   b. *[Un amico _] ha parlato di Gianni / che ho visto ieri
   A friend has spoken of G. / which (I) have seen yesterday
   c. *[Un libro _] è uscito di Chomsky / che è stato scritto da Chomsky
   A book appeared of C. / which was written by C.
   d. *[Un libro _] è stato pubblicato di Chomsky
   A book has been published of C.
   / which was written by C.

Better results (though slightly deviant for some speaker) are obtained with a post-verbal subject in Italian:

(13) a. È uscito [un libro _] ieri di Chomsky / che è stato scritto da C.
   (it) appeared a book yesterday of C. / which has been written by C.
   b. Ha salutato Gianni [un signore _] ieri che nessuno conosceva
   (pro) has greeted G. [a man] yesterday who nobody knew
   ‘A man who nobody knew has greeted G. yesterday’

On the other hand, EXT is fairly acceptable in Italian from a direct object (14a) or an indirect object (15a), but it yields grammatical results only when the interrupting constituent is “light” (prepositional modifiers seem to block EXT as shown in (14b) and (15b)):

(14) a. Gianni ha mangiato [un panino _] ieri con il prosciutto / che era avariato
   G. has eaten a sandwich yesterday with ham / which was rotten
   b. *G. ha mangiato [un panino _] in fretta con il prosciutto / che era avariato
   G. has eaten a sandwich in a hurry with ham / which was rotten
(15) a. Gianni ha mangiato un panino [con un’amica _] ieri
di suo fratello / che era ammalata
G. has eaten a sandwich with a friend yesterday
of his brother / who was sick
b. *Gianni ha mangiato un panino [con un’amica _] in fretta
di suo fratello / che era ammalata
G. has eaten a sandwich with a friend in a hurry
of his brother / who was sick

Notice that not only PP adjuncts (14b, 15b), but also the presence of an extra argument (16b) results in a degradation of EXT from direct object in Italian:

(16) a. Gianni ha spedito [una lettera _] ieri senza francobollo
G. sent a letter yesterday without stamp
b. ??Gianni ha spedito [una lettera _] a Maria senza francobollo
G. sent a letter to M. without stamp

Last consideration is related to “constituent-boundedness”: it seems that embedded constituents cannot be hosts of EXT in English (17a) / Italian (17b) while they can in Dutch (17c) (De Vries 1999)4.

(17) a. *[A review of [a book _]] appeared by three authors
b. *È apparsa [una recensione [di un libro _]] ieri di tre autori
   is appeared a review of a book yesterday by three authors
c. Ik heb [de papieren [van de man _]] gecontroleerd die een rode koffer
   I have [the papers of the man _] checked who a red suitcase
   droeg.
   carried

In sum:
i. both arguments and adjuncts can be host for EXT in English and Dutch;
ii. Italian marginally allows the last argument/adjunct to be host for EXT, and no “heavy” (prepositional) modifiers/arguments should intervene between the extraposed constituent and the related host;
iii. EXT from embedded constituents is generally impossible in English/Italian but it seems to be possible in Dutch;

2.3. Where
Looking at standard C-command tests, an extraposed element is not C-commanded by its host (Culicover and Rochemont 1997) since we get a principle C effect in the following examples5:

(18) a. I sent her, [many gifts _] last year that Maryi didn’t like.
b. *I sent her, [many gifts that Maryi didn’t like] last year.

4 But see English counterexamples provided by Strunk & Snider (2013) and the discussion in §5.8.
5 Following Larson (1988), the pronominal object C-commands the second object.
Do the “right” move  Cristiano Chesi

Looking at condition C bleeding under wh- movement and its sensitivity to the complements (19a) vs. adjuncts (19b) distinction (Van Riemsdijk and Williams 1981, Lebeaux 1988) so does EXT ((19c) vs. (19d), Taraldsen 1981, Fox & Nissenbaum 1999)) we could conclude that the object “reconstructs” under the host while the adjunct does not:

(19)

a. ??/*[Which book about John’s library] did he read _?
b. [Which book from John’s library] did he read _?c. ??/* I gave him [a picture of John’s mother] yesterday.
d. ??/* I gave him [a picture from John’s collection] yesterday.

As for the attachment point, it has been proposed that the constituents extraposed from subjects attach to IP (availability of stranding if the VP is elided (20a,b)), while the constituents extraposed from objects attach to VP (no stranding option if the VP is elided (20c-d)), i.e. an extraposed phrase is adjoined to the first maximal projection that dominates the phrase in which it originates (Baltin 1981, 2006). This is sufficient also to account for the classical nested dependency, (21a) vs. (21b):

(20)

a. Although [IP not [many people _] would [VP ride with Fred] who knew just him], some [IP would [VP _] who knew his brother].
b. Although [IP [no reviews _] [VP appeared] of Chomsky’s book],
one [IP did [VP _] of Jakobson’s book].
c. *Although he didn’t [VP call [people _] up who are from Boston],
he did [VP _] who are from New York.
d. *Although he didn’t [VP call [people _] up from Boston], he did
[VP _] from New York.

(21)

a. [IP [Someone _1] [VP picked [some books _2] up [which were lying on the table]2 VP] [who really didn’t want to]1 IP]
b. *[IP [Someone _1] [VP picked [some books _2] up [who really didn’t want to]1 VP]
[which were lying on the table]2 IP]

Such constraints on the attachment site have been captured by Williams’ generalization (1974):

(22) Williams’ generalization
When an adjunct β is extraposed from a “source NP” α, the scope of α is at least as high as the attachment site of β (the Extraposition site).

3. Another case of rightward movement: (Heavy) NP-Shift
Following Ross’ (1967) argument, reordering could also be considered as an instance of rightward movement; the direct object crossing over the indirect one or an adjunct is an example of this construction (Staub et al. 2006):
Lucy ate _ with a fork [the extremely delicious, bright green broccoli]

This “rightward movement analysis” has been seriously challenged by Larson (1988) and Kayne (1994), somehow reformulated in Jackendoff (1990) and integrated by Belletti and Shlonsky (1995). In the following pages, I will try to highlight some interesting parallelisms/differences that relate Heavy NP-Shift (henceforth HNPS) to EXT, suggesting that the rightward movement analysis might be on the right track in some specific case.

3.1. What
Looking at HNPS in English, we can observe that both direct and prepositional objects can undergo the alleged shifting operation:

(24)

a. I gave [the books which my uncle left to me as part of his inheritance] to Bill

b. I gave _ to Bill [the books which my uncle left to me as part of his inheritance]

c. I talk _ all the time [to my uncle who left me an enormous inheritance]

Even though such operation seems to be optional, a fairly evident bias (Pinker 1994:131) is attested for preferring certain shifted versions, (25a), of “heavy” constituents among many logically possible alternatives, (25b):

(25)

a. In my laboratory we use it as an easily studied instance of mental grammar, allowing us [to document [in great detail] [the psychology of linguistic rules] [from infancy to old age] [in both normal and neurologically impaired people], [in much the same way [that biologists focus on the fruit fly Drosophila to study the machinery of the genes]]]

b. ?? In my laboratory we use it as an easily studied instance of mental grammar, allowing us [to document [in much the same way [that biologists focus on the fruit fly Drosophila to study the machinery of the genes]]] [in both normal and neurologically impaired people], [in great detail] [the psychology of linguistic rules] [from infancy to old age]]

In general, “light” NP-shift is not possible in English, (26a’), while, in Italian, “shifting” seems to be freely applicable in the very same context, (26b’):

(26)

a’. In my laboratory we use it as an easily studied instance of mental grammar, allowing us [to document [in great detail] [the psychology of linguistic rules] [from infancy to old age] [in both normal and neurologically impaired people], [in much the same way [that biologists focus on the fruit fly Drosophila to study the machinery of the genes]]] [in both normal and neurologically impaired people], [in great detail] [the psychology of linguistic rules] [from infancy to old age]]

In (25) the heavy constituent is in Italic.

However, different orders have different informational structure implications (Belletti and Shlonsky 1995, Zubizarreta 1998). See discussion in §3.2.

---

6 A “heavy” NP is usually considered a NP containing a sentential/relative clause (Ross 1967). This (insufficient) definition will be better discussed in §5. In (25) the heavy constituent is in Italic.

7 However, different orders have different informational structure implications (Belletti and Shlonsky 1995, Zubizarreta 1998). See discussion in §3.2.
Do the “right” move

Cristiano Chesi

(26)a. I gave the books to Bill
   a’. *I gave to Bill the books
   b. pro ho dato i libri a Bill
   b’. pro ho dato a Bill i libri

In brief:
   i. Shifting is an optional operation;
   ii. Shifting only targets “heavy” elements in English, but it can also target light constituents in Italian;
   iii. Shifting is the preferred option for “heavy” constituents.

3.2. Where and why
Evidences that the “shifting” operation is not (always) a uniform “rightward” movement come from Italian (Belletti and Shlonsky 1995): The clitic ne (literally “some of them”) can be extracted only from the base object position (27b), but not, for example, from a post-verbal subject of unergatives (28b), Belletti and Rizzi 1981, Burzio 1986). Ne cliticization with light NPs is only possible if they are not shifted, (29a) vs. (29b). However, with heavy (30a) and focussed (30b) NPs that appear in a “shifted” position, ne-cliticization is also possible. This has been interpreted as the signature of the fact that the “shifted” object, e.g. (30), is, indeed, in his base position:

(27) a. Ho letto molti libri.
   (I) read many books.
   b. Ne ho letti molti.
   (I) of-them read many.
(28) a. Hanno lavorato molti operai.
   have worked many workers.
   b. *?Ne hanno lavorato molti.
   of-them have worked many.
   (I) have given one/three to Gianni.
   b. *?Ne ho dato/dati a Gianni uno/tre.
   (I) have given to Gianni one/three.
(30) a. Ne ho dato/dati a Gianni uno/tre che mi avevano consigliato la settimana scorsa.
   (I) have given to Gianni one/three which to-me (they) have suggested last week
   ‘I gave to Gianni one/three which they suggested to me last week’
   b. Ne ho dato a Gianni uno solo.
   (I) have given to Gianni one only
   of-them (I) have given to Gianni one only

Because of this paradigm, Belletti and Shlonsky (1995) proposed that when the light object is postponed, it does not occupy anymore the object position (Light NP Postposing). This strategy is specific to Italian (and not available in English) because
of the availability of subject/object pro in this language and it results in a focalization of the shifted constituent (this reminds of the obligatory focalization of post-verbal subject, Calabrese 1982, (31a)). Assuming that the focalized position to the “right edge” of the VP is unique, this solution can capture the impossibility of “light”-shifting both of the subject and the object at the same time8 (31b):

(31) a. ¿Ha dado un libro a Maria Gianni.
    has given a book to Maria Gianni.
 b. *(?)Ha dado a Maria un libro Gianni.
    has given to Maria a book Gianni.

This solution, however, seems to be adequate only for the shifting of light NPs but not for the (adverbially) focalized object in (30b) for which Belletti and Shlonsky (1995) propose that the PP must scramble before the NP, which, in fact, stays in situ allowing ne extraction. This solution is supported by the behaviour of the weak dative pronominal element loro that is forced to move closer to the verb (Cardinaletti 1991, (32a)) as predicted in the PP scrambling hypothesis; in this case, ne extraction from the direct object is allowed as expected, (32b):

(32) a. ho dato loro un libro.
    (I) have given to-them one book.
 b. ne ho dato loro uno.
    of-them (I) have given to-them one.

Unfortunately, the scrambling analysis of PP does not explain why the shifted heavy-NP behaves as if it were in an A’-position (instead of an A-position according to the ne cliticization test) when it licenses parasitic gaps:

(33) a. I crossed _ without recognizing _ a classmate with whom I attended primary school.
    (I) have crossed without (to) recognize that classmate with whom avevo frequentato le scuole elementari.
    (I) have attended the primary schools.
 b. *Ho incrociato _ senza riconoscere _ quel compagno con cui
    (I) crossed without to recognize that classmate.
    Ho incrociato senza riconoscere _ quel compagno.

The contrast between (33b) (heavy NP) and (33c) (light NP shifting) shows that heaviness is in fact the crucial property that triggers, in a cross-linguistic uniform way, a displacement operation that removes the direct object from its base position. This is the classic understanding of HNPS.

Before concluding this paragraph, I wish to stress a symmetry between HNPS and movement, depending on the source position in double object constructions:

---

8 Italian post-verbal subjects are more restricted than in other null subject languages such as Spanish (Belletti 2001:70f.).
extractions from the first object (34a,b) seem to be much more difficult than from the second one ((34c,d), Culicover and Rochemont 1997:15):

(34)  a. *Bill [gave _ ] the book yesterday anyone who wanted it
    b. *Who did Bill [give t₁ ] the book yesterday
    c. Bill [gave [John ] _ ] yesterday the book that he was looking for
    d. What did Bill [gave [John ] t₁ ] yesterday

Summarizing the main facts discussed in this section:
  i. Argument re-ordering is cross-linguistically sensitive at least to two factors: heaviness (classic NP-shifting), focalization of light constituents and availability of empty pronominal elements (to satisfy substantive criterial and/or verbal theta-role requirements when the object is not in its canonical base position, i.e. next to the verb);
  ii. Movement and HNPS behave the same (both degraded) with respect to extraction from the first object in double object constructions.

4. Some of the solutions proposed and their problems

In this section I will briefly review the main analyses that have been proposed to account for the data described in §2 (and partly in §3, which have been already discussed in §3.2). As it will be clear, none of them results in a sufficient coverage of the empirical evidence reported, especially in accounting for the three properties discussed in §1 (clause-boundedness, definiteness constraint and adjunct/argument distinction insensitivities to EXT).

4.1. Analysis 1: Classic rightward movement

The first solution that has been proposed to account for EXT (that is straightforwardly suitable also for HNPS) has been a rewriting transformation (Rosenbaum 1965:11), adapted by Ross (1967) as follows:

(35) Extraposition from NP (Ross 1967:4):

\[
\begin{array}{ccc}
 X & [NP - S] & Y \\
 1 & 2 & 3 \\
 \hline
 \rightarrow & \text{OPTIONAL} & 1, \emptyset, 3+2
\end{array}
\]

According to this rule, the underlying structure of (10a) is (10b), repeated below:

(10) a. [A book _ ] appeared which was written by Chomsky
    b. [A book which was written by Chomsky] appeared

It is easy to show that this rule is too powerful and it cannot block ungrammatical transformations like (36b):

(36) a. [S [NP [S That [NP a gun [S which I had cleaned]] [VP went off]] ]] [VP surprised no one]]
    b. *[S [NP [S That [NP a gun _ ]] [VP went off]] ]] [VP surprised no one] [S which I had cleaned]]
For this reason Ross proposes an ad-hoc constraint, that excludes (36b) forcing clause-boundedness:

(37) **Right Roof Constraint** (RRC, Ross 1967)

An element cannot move rightward out of the clause in which it originates.

At least three things remain unexplained under this fairly simple analysis:

a. Why is movement much more constrained to the right than to the left?

b. What does trigger Extraposition?

c. How can we account for the clause-boundedness of the rule?

4.2. Analysis 2: Base generation

The movement analysis has been criticized because of an incomprehensible asymmetry with respect to standard leftward movement (Akmajian 1975): why should wh-movement escape subjacency by successive cyclic applications of the movement operation while EXT cannot? Also from a licensing perspective, the movement analysis poses some problems: for instance, when a rRC is extraposed from a subject position, the trace it leaves would violate the Empty Category Principle (Chomsky 1981). To solve these problems an alternative analysis has been proposed, among others, by Guéron and May (1984), Culicover and Rochemont (1990-97); this analysis relies on the alleged satisfaction of the Complement Principle at some point in the derivation:

(38) **Complement Principle** (Guéron and May 1984)

In a sequence of categories \( \alpha_i \beta^i_1 \ldots \beta^n_i \) in a structure \( \Sigma \beta^i_1 \ldots \beta^n_i \) are complements to \( \alpha_i \) only if \( \alpha_i \) governs \( \beta^n_i \) (where \( \alpha \) governs \( \beta \) iff \( \alpha \) and \( \beta \) are dominated by all the same maximal projections, and there are no maximal projection boundaries between \( \alpha \) and \( \beta \)).

This principle obviously cannot apply at the surface structure, where the relevant locality (government) requirements between the host and the extraposed constituent are not met. They propose, then, that the principle must be satisfied at LF, crucially after Quantifier Raising (QR). This intuition elegantly accounts for the “locality” (clause-bounded because of QR) and for the definiteness constraint (no quantification, no QR) of EXT\(^9\), but it contains at least one important flaw: the complement principle does not make any distinction between arguments and adjuncts (in both cases, no maximal projection boundaries, i.e. CP, intervene between the noun/verb and the DP/PP, then, according to (38), in both cases, arguments and adjuncts are governed), but this seems to be a very productive distinction as shown in (19).

---

\(^9\) Walker (2013), running an experiment to test the speaker performance with respect to the Definiteness Constraint, verifies that this seems to be only a weak constraint, i.e. it can be violated in some significant way. I would prefer sticking on the strongest assumption and accept the performance issues raised in her study as the result of the availability, for some speaker, of a null existential Q operator on the otherwise simply definite DP (e.g. “(∃) the girl appeared...”). Notice that the verb of appearance facilitate this interpretation, thus also the contrast with respect to other verb classes seems less mysterious.
4.3. Analysis 3: Modification based account

Kayne (1994) proposes that the host and the extraposed constituent are generated together, then the host moves (to a C-commanding position) while the extraposed element is stranded in its base position.

Such analysis is the only possible solution assuming the Linear Correspondence Axiom (and its implicit ban on rightward movement \(^{10}\)). This proposal easily accounts for the Definiteness Constraint (Diesing 1992, (3)) since non-constituents cannot be moved; given the relative clause structure and the DP structure proposed by Kayne, this analysis is pretty straightforward (“the book” in (39a) is not a constituent, so it cannot move, while “two/a/those book(s)” is a constituent, so it can move stranding the RC behind as expected):

\[
\begin{align*}
(39) \ a. & \ [\text{DP[D the]}][\text{CP[NP book]}][\text{C[C that][TP I [VP[V read][NP ti]]]}] \\
& \ b. \ [\text{DP[D 0]}][\text{CP[NP two/a/those book(s)]}[\text{C[C that][TP I [VP[V read][NP ti]]}]]
\end{align*}
\]

Under this perspective the Right-Roof Constraint (37), follows from theta-role requirements and/or LCA:

\[
(40) \ *\text{The fact that [somebody], walked into the room is irrelevant [ t, who I knew]. (Kayne 1994:118) }
\]

Nevertheless, this account presents some problems:

a. Extraposed constituents seem not to be C-commanded by their host (e.g. (18));

b. Examples like (10c) remain unexplained (since [P NP] should not be considered a constituent using Kayne’s analysis);

c. The stranding analysis would not easily predict the nesting dependencies reported in (21) (because of the order of complements, i.e. [[V [DPi [EXT1]] [DP2 [EXT2]]]], unless we assume some sort of scrambling, standard movement operations would affect first the inner constituent, i.e. [[DP2 ] [V [DPi [EXT1]] [t2 [EXT2]]]], then the higher one, leading to a cross-serial dependency, i.e. .[[DP1 ] [[DP2 ] [V [t1 [EXT1]] [t2 [EXT2]]]])].

4.4. Analysis 4: “Mixed” account

The last proposal I would like to review is the solution proposed by Fox and Nissenbaum (1999). The general idea is that “overt” and “covert” movement simply differ for the pronunciation of the head of the chain (overt) vs. pronunciation of the tail of the chain (covert). The relevant covert operation, in this case, as in Guéron and May’s (1984) analysis, is QR; what is crucial in this proposal is the assumption that covert operations, i.e. QR, can precede overt ones: EXT, then, can be the result of QR plus late merge of an adjunct. This is the schematic derivation Fox and Nissenbaum propose:

\[
(41) \ a. \ [\text{CP John [VP [VP picked [DP a book] up] [DP a book]]]}
\]

\[
(41) \ b. \ [\text{CP John [VP [VP picked [DP a book] up] [DP a book]]] (QR)}
\]

\(^{10}\) If \(A\) asymmetrically C-commands \(B\), then terminals dominated by \(B\) can not precede terminals dominated by \(A\) (Kayne 1994:33): this is why rightward adjunction could not be linearized.
iii. \([\text{CP John} [\text{VP picked [DP a book] up}] [\text{DP [DP a book] [CP which he really enjoyed]]}]\) (late merge)
iv. \([\text{CP John} [\text{VP picked [DP a book] up}] [\text{DP [DP a book] [CP which he really enjoyed]]}]\) (PF deletion of the head of the QRed nominal chain)

As for EXT of complements, if we assume that thematic requirements need to be satisfied everywhere (both at Deep Structure and at LF as predicted by the Complement Principle), the complements need to be merged before, then QRed afterward; so, no late merge of complements is available.

It is somehow surprising that the rightward (not leftward, as usually assumed) QR operation is not discussed in any detail in their work; nonetheless, this analysis allows us to capture many interesting facts:

1. The clause-boundedness of Extraposition is determined by the clause-boundedness of QR (as in Guéron and May (1984), Culicover and Rochemont (1990-97));
2. The definiteness constraint is readily captured: definite NPs cannot be QRed (as in Guéron and May (1984), Culicover and Rochemont (1990-97));
3. Late merge of adjuncts vs. complements captures the condition C bleeding effect discussed in (19).

Moreover, the late merge hypothesis would correctly predict that extraposed RCs do not reconstruct (Wilder 1995):

\[(42)\]
\begin{align*}
a. \text{We talked [about her claim ] yesterday that Mary will hire Peter.} \\
b. \text{I gave him [an argument ] yesterday that supports John's theory.}
\end{align*}

Despite the coverage of this elegant and relatively simple analysis, one main question remains unanswered: why is QR assumed to be rightward?

5. The proposal: changing the orientation of the derivation

In order to account for the set of characterizing properties of EXT and HNPS, I propose to adopt a radically different minimalist derivational perspective: constituents must be built not from bottom to top and, consequently, from right to left (at least in right branching languages) as generally assumed, but strictly from left to right, expanding lexical “expectations” in a principled, top-down, way.

5.1. Deriving phrase Top-down

A top-down (Chesi 2004-12) derivation is a generative procedure that assembles lexical items in phrase structures, using structure building operations inspired by the Minimalist Program (Chomsky 1995-2008, Stabler 1997). The Phase-based Minimalist Grammar (PMG, Chesi 2007) that implements such (non-conventional) derivational direction requires a fully explicit lexicon, the formalization of the necessary structure building operations, and a clear specification of their domain of application.

As for the lexicon, I will adopt a simplified version of Stabler’s (1997) formalism: a lexical item is a ordered feature structure\(^{11}\) composed by functional features (prefixed

---

\(^{11}\) Here to express ordered sets I will use squared brackets: e. g. \([A, B, C]\) or simply \([A B C]\).
by the “+” diacritic, expressing functional properties/positions, e.g. +D(eterminer)),
selection features (characters prefixed by the “=” diacritic, expressing thematic
requirements in terms of necessary additional constituents that must be present in the
structure in order for the expression to be grammatical), categorial features
(unprefixed characters expressing N(ominal), V(eral) and A(djectival/dverbial)
heads) and, in the end, phonetic/semantic features (that I do not discuss here and I
indicate using placeholders). This is an example of lexical item (as in “John
runs”)\textsuperscript{12}:

\begin{equation}
(43) \quad \begin{aligned}
&\text{features:} &\text{functional} &\text{selectional} &\text{categorial} &\text{phon/sem} \\
&\quad [+T_{pres} +Agr_{3-sing} =+[+D N] V \text{ runs}] 
\end{aligned}
\end{equation}

Structure building operations will target lexical items and phrases and access their
feature structure in a constrained way. To understand such constraints we distinguish
operations that establish local relations from the operations creating non-local ones.
Looking at local relations, if we assume that:
\begin{itemize}
\item[i.] Merge is responsible for local phrase structure building,
\item[ii.] it is binary,
\item[iii.] it is successful between A and B (with A and B either lexical or phrasal) if
\quad and only if A selects B or B selects A\textsuperscript{13},
\item[iv.] the item that selects also projects,
\end{itemize}
a natural Top-Down constraint follows:

\begin{equation}
(44) \quad \begin{aligned}
\text{The selecting item is computed before the selected one.}
\end{aligned}
\end{equation}

This is so, because of the nature of the hierarchical structure: if A selects B, as in (45),
A will be computed before B if the tree is explored Top-Down. I propose, then, that
computing A will generate the expectation of B, because of the select feature on A.

\begin{equation}
(45) \quad \begin{aligned}
&\begin{array}{c}
\quad [\Rightarrow A] \\
\quad [\Rightarrow A] \\
\quad [\Rightarrow B]
\end{array}
\end{aligned}
\end{equation}

Pushing forward this idea, the Top-Down Merge reduces to local lexical selection
satisfaction (and, eventually, unification, Shieber 1985, Chesi 2012:51).

\textsuperscript{12} Let us take this as a simplified example. Obviously we could decompose such lexical items in its
morphological subparts ([=+[+D N] V run] [+T_{pres} \emptyset] [+Agr_{3-sing} s]) but this would require extra
machinery to be discussed. Similar considerations hold for feature structures: in this paper, I assume a
pure privative system, i.e. any feature is atomic and its presence/absence within a lexical item is
completely arbitrary (Adger 2007). In this sense, every lexical item is an “exception” (Chomsky
1995:235). In many cases, using full feature structures could make many tasks much more elegant

\textsuperscript{13} Stabler (1997), for instance, uses this restriction to limit Chomsky’s (1995) Merge. Similarly,
Collins (2002) proposes that Merge itself is triggered by saturation considerations. The recent
discussion on ‘labeling’ of tree nodes seems to me to go exactly in this direction (Cecchetto & Donati
2010): the ‘label’ (i.e. the interpretable result of Merge, accessible to other Merge operations) is the
‘probe’, i.e. the merged item that selects the other
As for non-local relations, some important constraint might follow from an answer to the question “How many features might be expected (hence selected) at once?”. My tentative answer diverges from the minimalist one (that is, just one categorial feature, e.g. \([-\text{N D}]\)), and it is crucially related to the minimalist notion of Phase. From a Top-Down perspective, the features that can be selected or expanded correspond to the minimal computational domain within which a non-local dependency should be evaluated. By assumption, this domain matches with the extended projection (Grimshaw 1991) of a lexical head (N, V or A) and its selected arguments. In this sense, the generally accepted structures in (46)a and (46)b will be replaced by (46)a' and (46)b' (i.e. extended projections of V and N, respectively):

\[
\begin{align*}
\text{(46) a. } & & \text{b. } \\
\text{CP} & & \text{DP} \\
\text{[\(-\text{TP C}\)]} & & \text{[\(-\text{NP D}\)]} \\
\text{TP} & & \text{NP} \\
\text{[\(-\text{VP T}\)]} & & \text{VP} \\
\text{VP} & & \text{NP} \\
\text{a'. } & & \text{b'. } \\
\text{V} & & \text{N} \\
\text{+C} & & \text{+D} \\
\text{+T} & & \text{N} \\
\text{V} & & \text{V}
\end{align*}
\]

By using extended projection, selecting a DP (i.e. =DP) equals to say that a determined (+D) nominal (N) extended projection is expected (i.e. (46)b') while selecting a tensed declarative sentence would correspond to the expectation of the structure in (46)a'. In fact, because of the necessity to express the subject of predication, SVO languages should also include, a +S (criterial, in the sense of Rizzi 1997) feature in the verbal extended projection corresponding to a declarative sentence expectation, that is: [+C +S +T V].

Processing a “phase” means lexicalizing any expected feature using a compatible item taken from the lexicon. We predict that both phonologically null items ([+C \(\Omega\)] to indicate a declarative force, and [+S\(_{\text{nom}}\) \(\Omega\)] working as an implicit case marker) and features clusters ([+C\(_{\text{wh}}\) +D N what] or [+S +D N]) are present in our lexical knowledge. So it can happen that either lexicalized (e.g. [+C\(_{\text{wh}}\) +D N what]) or unlexicalized (e.g. [+S +D N]) feature clusters are used to expand just one expected features. By assumption, the “unexpected” features that are introduced must be “stored in memory” until they will be properly selected.

This is what triggers movement, i.e. a non-local dependency of the filler-gap, filler-first kind (Fodor 1978), both in the case of criterial wh- items and in the case of the pre-verbal subject; in an informal way this is how the derivation unrolls in the simple case of a finite declarative sentence, in a SVO language like English:
(47) a. a declarative sentence is expected: [+C +S +T V];
b. +C is lexicalized by a null [+C Ø];
c. [ [+Snom Ø] +D N] lexicalizes the +S expectation;
d. [+D N John] lexicalizes both +D and N, but since these features where unexpected, [+D N (John)] is stored in a memory buffer14;
e. [+T V =DP sings] lexicalizes +T and V at once;
f. =DP selection of [=DP sings] creates the landing site for an item compatible with [+D N] expectation;
g. [+D N (John)] is discharged from the memory buffer;
h. The derivation ends since no more features must be processed.

In sum, the Phase idea, and the Merge and Move operations are reformulated according to the logical constraints here discussed and produce a left-right derivation as expected both in parsing and generation15:

(48) Phase: it is the minimal computational domains including an extended projection of an expected lexical head and its selection requirements;
(49) Merge: it lexicalizes/expands the expected features (by unification, Shieber 1986) introducing in the left-most unlexicalized position, a compatible item taken from the lexicon;
(50) Move: it creates Long Distance (Filler-Gap, Filler-First) Dependencies by means of a Memory Stack; constituents are “moved” in the memory buffer when they introduce unselected (i.e. unexpected) features and retrieved/re-merged in the structure as soon as possible when properly selected (Moved items preempts the insertion, i.e. Merge, of new items from the lexicon).

5.2. Merge, Movement and Phase projection at work: nesting and the special status of the last selected argument

There are some logical possibilities on the order of application of the three operations just discussed that crucially result in different derivations/structures16. I assume that in the default case, we first apply Phase Projection, then Merge and in the end Move. Below, a sample derivation that shows how Phase Projection (PhP) and Merge operate in the case the Move buffer be empty17 and a base verbal phase is expected:

14 The null hypothesis is that features already used will not be re-merged twice; e.g. the first copy of a spelt out item is the one consuming its phonological features, this is why (John) is indicated under brackets. It is plausible to expect cross-linguistic parameterization on this, but I will not discuss these aspects here.
15 See Chesi 2007, 2012 for the full formalism and for some arguments on its generative power. This formalization (Phase-based Minimalist Grammars, PMGs) is a modification of Stabler’s 1997 (and its revision in Collins and Stabler 2011) proposal (Minimalist Grammars, MGs).
16 Interleaving in a way or another structure building operations results in grammars with different generative power (see Chesi 2007 for discussion).
17 I will use squared brackets to mark feature structures/constituency, but remember that features within the feature structure are ordered; =DP/=PP are meant to be shortcuts for =+[D N] and =+[K +D N] respectively, where +K is a case feature.
(51) Lexicon: $\{[-\text{DP} \rightarrow \text{DP} \rightarrow \text{PP} \ V \ gives], [+K \ to], [+D \ N \ John], [+D \ N \ children], [+D \ N \ candies]\}$

1. $\text{PhP}(\sim V) \rightarrow \{V\}$ (this is a default phase projection that simply instantiates a verbal phase)

2. $\text{Merge} \ (\{V\}, [-\text{DP} \rightarrow \text{DP} \rightarrow \text{PP} \ V \ gives]) \rightarrow [-\text{DP} \rightarrow \text{DP} \ V \ gives]$  

3. $\text{PhP}([-\text{DP} \rightarrow \text{DP} \rightarrow \text{PP} \ V \ gives]) \rightarrow [-\text{DP} \rightarrow \text{PP} \ V \ gives [+D \ N]]$  

4. $\text{Merge} \ ([-\text{DP} \rightarrow \text{PP} \ V \ gives [+D \ N]], [+D \ N \ John]) \rightarrow \ [\text{DP} \rightarrow \text{PP} \ V \ gives [John]]$  

5. $\text{PhP}([-\text{DP} \rightarrow \text{PP} \ V \ gives [John]]) \rightarrow [\text{V} \ gives [John] [-\text{PP} \ (gives) [+D \ N]]]$  

6. $\text{Merge}(\{\text{V} \ gives [John] [-\text{PP} \ (gives) [+D \ N]], [+D \ N \ candies]\}, [+D \ N \ candies]) \rightarrow \ [\text{V} \ gives [John] [-\text{PP} \ (gives) [+D \ N \ candies]]]$  

7. $\text{PhP}(\{\text{V} \ gives [John] [-\text{PP} \ (gives) [+D \ N \ candies]]\}) \rightarrow \ [\text{V} \ gives [John] [-\text{PP} \ (gives) [+D \ N \ candies] [+K \ +D \ N]]]$  

8. $\text{Merge} \ (\{\text{V} \ gives [John] [+K \ +D \ N], [+K \ +D \ N \ to \ children]\}) \rightarrow \ [\text{V} \ gives [John] [+K \ +D \ N \ to \ children]]$  

4. $V$  
6. $V$  
8. $V$

In this derivation, $\text{Phase Projection}$ creates an empty (i.e. unlexicalized) constituent to the right edge of the structure depending on the first select feature of the processed lexical item; $\text{Merge}$ fills this position either with the most prominent element in the memory buffer (if any, and this is not the case in the example) or with an element from the lexicon (as in (51) where the memory buffer is empty).
In the case the memory buffer is not empty, storage and retrieval is constrained by phase. The reason for using phases follows from cognitive and computational considerations, that is, we want to keep “working memory” as small as possible (coherently with processing evidence, Baddeley et al. 2009). In this sense, I will retain the minimalist intuition that, at some point, the phase is “shipped-out”, and no further operations can tamper the set of dominance/precedence relations created up to this point. In this vein, we decided (Chesi 2004, Bianchi and Chesi 2008) to fix this “memory limit” at the last operation triggered by the phase head, that is, the Phase Projection of the last selected argument. This produces important computational advantages (recasting the true vs. tail recursion distinction, Abelson and Sussman 1996, in phasal terms, Chesi 2004-12) and allows us to make a crucial distinction between the right recursive branch of the tree, i.e. the last selected complement (the sequential phase, Bianchi and Chesi 2008), that once projected close the previous phase (since this allows for unlimited recursion, i.e. =X2 in (52)), and the nested phases (Bianchi and Chesi 2008), i.e. unselected phases resulting from expansion of functional features or phases that are not the last selected one (that, in case of recursion, would create centre-embedding, Bever 1970):

(52)

Posing the constraint that every phase has its own memory buffer and that the inheritance of the content among memory buffers is sensitive to the sequential/nested distinction, we can constraint, in an empirically tenable way, the usage of the memory buffers and, as a consequence, the movement operation:

(53) Constraint on the memory buffer inheritance
The content of the memory buffer is either integrated within the phase, or discharged in (the memory buffer of) the last selected phase.

(54) Success condition
Memory buffers must be empty at the end of the computation.

A simple example showing how movement operates is subject movement in SVO languages. According to what I said, we expect the basic word order of transitive verbal phrases to be VSO, as in (51). In order to derive the relevant SVO language we would have two options:

i. the verbal head could be spelled out at the lower VP-shell;
ii. the subject is generated first in preverbal position, then moved.

While the first solution could be cross-linguistically promising, here I propose that SVO order is obtained by exploiting the second strategy. This is a graphical
representation of the sketched derivation of a simple SVO sentence in English (the numbers labelling the arrows indicate the sequential order in which the operations apply).

\[(55)\]

In the following pages, I will use a box-notation\(^{18}\) to visualize derivations/structures as the one in (55). Nested phases are expressed by forefront boxes; when phases are unselected (e.g. adjuncts, like “at the market” in (56b) or the preverbal “John” in (56b)) they are indicated with grey boxes; in case of selected arguments (white boxes) I expect only the last selected argument that terminates the phase to qualify as sequential (e.g. “the apples” in (56a)), then in other cases, the arguments should be nested (as in the case of “the apples” in (56b), since Ph\(_3\) is followed by a PP modifier, Ph\(_4\), related to the matrix phase) though we might expect some degree of cross-linguistic variation in this respect. In the sequential case the boxes are white and aligned as in (56a):

\[(56)\]

These assumptions are sufficient to capture (successive cyclic, (57b)) movement, (57a), predicting the correct locality constraints\(^{19}\):

\(^{18}\) The “box-notation” is due to Valentina Bianchi. This notation allows us to keep track of the derivation in a compact and meaningful way. In a nutshell: squared boxes are phases, rounded boxes are memory buffers (the elements are ordered from left to right, i.e. the rightmost one is the first one to be remerged), indexes mark the univocal phase-memory buffer dependency.

\(^{19}\) In the original proposal, Chesi 2004, the memory buffer was structured in slots in order to capture Relativized Minimality effects (Rizzi 1990). Under the cue-base retrieval assumption (Van Dyke & McElree 2006) this hypothesis seems unnecessary.
Do the “right” move

Cristiano Chesi

In (57a), the steps of the derivations are the following ones:

1. A default verbal phase is projected (P1):
   \[ \text{PhP}(\text{Wh-question}) = [+\text{wh} +T +S V]^{20} \]

2. Since the verbal phase is interrogative, this functional feature has to be explicitly marked; in English this is done by merging the relevant wh-element within the specific “criterial” position. This is how \([+\text{wh} +D N \text{what}], \text{phase P2} \) (computed as a nested phase), is introduced in the derivation:
   \[ \text{Merge}([+\text{wh} +T +S V], [+\text{wh} +D N \text{what}]) = [+\text{wh}[+\text{wh} +D N \text{what}] +T +S V] \]

3. Since \([+D N]\) are unselected \([+D N (\text{what})]\) is inserted (step 1) in the memory buffer (M1) of the matrix V-phase (P1):
   \[ \text{Move}([\text{what}]) = M1< [+\text{wh} +D N (\text{what})] > \]

4. did is compatible (unifiable) with a tense functional specification of the matrix V-phase, then licensed in this position:
   \[ \text{Merge}([+\text{wh} +T +S V], [+T \text{did}]) = [+\text{wh} +T[+T \text{did}] +S V] \]

5. \([+D N \text{John}]\) (phase P3, again computed as a nested phase) is introduced to satisfy a subject-criterial (in the sense of Rizzi 2006) requirement (functional specification of P1, i.e. +S, triggering \([+S[\emptyset] +D N]\) phase expansion) and moved in the memory buffer since \([+D N]\) are unselected (step 2):
   \[ \text{Move}([+\text{wh} +T did] [+T \text{did}]) = M1< [+\text{wh} +T did] +S[+S[\emptyset]] +D N \text{John}] > \]

---

20 It is fair to assume that subject-aux inversion is decided (as parameterized option) at this level. PhP stands for the Phase Projection function.
6. \([-[\text{+D N]} \text{+D N}] \text{买} \) is merged as the head of the matrix V-phase \((P_1)\). Since it has two selection requirements to be satisfied (an agent and a patient, both N-phases), these select features will project two phases, \(P_4\) and \(P_5\), in sequence:

\[
\text{Merge}(\text{[+wh [+D N \text{what}]] [+T \text{did}]] [+S [\text{[+D N \text{John}]] \text{买}]], \text{[-[\text{+D N]} [\text{+D N} \text{买}]]) =}
\text{[+wh [+D N \text{what}]] [+T \text{did}]] [+S [\text{[+D N \text{John}]] [\text{买}]]]}
\]

\[
\text{PhP}(\text{[-[D N]}]) = \text{[+wh [+D N \text{what}]] [+T \text{did}]] [+S [\text{[+D N \text{John}]] [\text{买}]]]}
\]

7. \(P_4\) is a nested phase and will be unified by (re-)merging \(P_3\) (the first accessible element in \(M_1\));

\[
\text{Merge}(... \text{[[-[\text{+D N]} \text{买} [\text{[-[\text{+D N]} \text{买}]]) =}
\text{[-[\text{+D N]} \text{买} [\text{[-S] +D N (\text{John})]) ]]}
\]

8. \(P_5\) is the last Phase Projection: it is a selected/sequential phase where the last selectional requirement of the previous phase will be lexicalized by merging \(P_2\):

\[
\text{PhP}(\text{[-[D N]}]) = \text{[[-S] +D N (\text{John})][\text{买} [\text{[-[D N]}])]]}
\]

\[
\text{Merge}(... \text{[[-[\text{+D N]} \text{买} [\text{[-[\text{+D N]} \text{买}]]) =}
\text{[-[\text{+D N]} \text{买} [\text{[-S] +D N (\text{what})]]) ]]}
\]

9. The derivation here terminates since no more selection features must be computed and the memory buffer is empty as required by \((54)\).

Reiterating these steps recursively we can easily derive the successive cyclic movement in \((57b)\).

5.3. Right hand adjuncts

As in any antisymmetric framework, which with the present approach share the necessity of deriving linear order from hierarchical one, we would not expect adjuncts, which are functional-related positions (i.e. functional feature in the extended projection of a lexical head in the current approach, specifiers in an antisymmetric approach), to be placed to the right of the head. To explain the fact that certain adjuncts sometimes are rightward, we need either to rely on an interleaving option between Phase Projection and Merge, or on the possibility of stacking the Phase Projections after the last selected complement. Before discussing which solution better fits with the relevant data under analysis, let us understand how we can determine different head-adjunct orders. I assume that a rightward “shifting” operation becomes an option when a functional feature is “complex”, namely when it bears a select feature. In this case, two possibilities should be available (delayed Phase Projection or Stacking) and we would obtain \((58b)\) from \((58a)\) (notice the minimal difference between an adjunct selected by a functional feature \((58a)\) and a selected (prepositional) argument \((58c)\)).

\[
\text{(58) a. b. c.}
\]

\[
\text{[\text{+S} +D N (\text{John})]][\text{买} [\text{[-[S] +D N (\text{what})]])]}
\]

\[
\text{[\text{[-[S] +D N (\text{John})][\text{买} [\text{[-[S] +D N (\text{what})]])]}}
\]

\[
\text{[\text{[-[S] +D N (\text{John})][\text{买} [\text{[-[S] +D N (\text{what})]])]]}
\]

127
There are reasons to believe that such solution is not arbitrary and it is justified by complexity reduction strategies (Chesi 2007) aiming at “minimizing nesting” as long as the relevant functional feature is not criterial and the dependency between the selecting feature/position and the “shifted” element is recoverable. On this line, we can state a general principle that should allow us to predict shifting of “complex” constituents:

\[(59) \text{Minimize Nesting}\]

When a nested phase bears select features (i.e. it is heavy/complex), it would rather be (partially) processed in a phase-peripheral position (e.g. to the right\(^{21}\)), unless licensed functional/thematic features be unrecoverable.

“(Partially) Processed in a phase-peripheral position” means that the whole constituent could be shifted (unless functional/thematic features be unrecoverable) or else, only the selected phase(s) be projected in a phase peripheral position. The definition of Phase Projection however does not allow us to project freely under non-local conditions, therefore the phase head has to be somehow moved in the new peripheral position. The next paragraph suggests a trigger for this movement operation.

As for the two available solutions, they are clearly distinct as shown below:

\[(60)\]

\[(\text{a}) \text{ Delayed Phase Projection} \]

\[
[[+A \rightarrow (X)] [+B \rightarrow (Y)] C \ldots] \rightarrow [[+A \rightarrow (X)] [+B \rightarrow (Y)] C \ldots X[A] Y[B]]
\]

\[(\text{b}) \text{ Stacking} \]

\[
[[+A \rightarrow (X)] [+B \rightarrow (Y)] C \ldots] \rightarrow [[+A \rightarrow (X)] [+B \rightarrow (Y)] C \ldots Y[B]] X[A]
\]

The mirrored prediction made by stacking seems to be more promising (remember the nesting dependency discussed in (21)) and it will be considered first.

5.4. Rightward Quantifier Raising

In a left-to-right derivation there is no room for a leftward movement in the classical sense. Quantifier Raising is not an exception in this respect. What we proposed (Bianchi and Chesi 2010) is that the long distance dependency through memory buffers fits with QR as well\(^{22}\). There are obviously differences with respect to the Move operation that we need to capture:

i. QR is covert (in the sense that the head of the chain is unpronounced);

ii. It is not feature-driven\(^{23}\);

iii. It cannot be freely cyclic\(^{24}\).

All these properties can be accounted for, if we assume that QR is an operation that computes a selected position and removes from it the quantified element (QP) which is not interpreted there; this element should be remerged (then interpreted) after the

\(^{21}\) This might be another possible parameter: head initial languages marginalize at the right edge (shifting), while head-final languages marginalize at the left edge of the matrix phase (scrambling).

\(^{22}\) Cf. lowering discussed in McCawley (1999).

\(^{23}\) But see Begelli and Stowell 1981 and discussion in Bianchi and Chesi 2010.

\(^{24}\) But see Cecchetto 2004 and discussion in Bianchi and Chesi 2010.
relevant nuclear scope has been created. It follows that QR is actually a rightward movement (as proposed, without discussion, in Fox and Nissenbaum 1999).

More precisely, we proposed that QR:

i. stores a QP in a dedicated memory buffer of the current phase (Schlenker 2005);
ii. integrates a coindexed variable in the corresponding argument position or memory buffer (depending on where the QP is processed: selected vs. nested position);
iii. when the top-down computation of the current phase is concluded, the QP function is retrieved from the Q-buffer and takes scope over the structure built so far. The elements retrieved from memory buffers are (typically) not spelled out (footnote 14, Chesi 2004); hence QR is “covert”.

For example, in the computation of (61), P₁ (a verbal phase) is the first phase computed and the subject QP₁ (a quantified nominal phase) constitutes a nested phase. This is computed while the nuclear scope (P₁+P₂) is still incomplete. Then, the nested QP₁ is stored in the Q-buffer (Q₁) of the containing phase (P₁) while the coindexed variable is stored in the already discussed Move Buffer (M₁) and behaves as explained in §5.2.

Then we process the sequential phase (P₂) which is still part of the nuclear scope and only at the end of it, we can remerge the QP.

(61)

Notice that this is sufficient to explain clause-boundedness (independently predicting the Right Roof Constraint in (37)): since buffers are phase-local, a QP cannot attach to a superordinate phase, this explains the impossibility of getting inverse scope in sentences like the following one:

(62) Someone thinks [CP that every man loves Mary]. (∃>∀; *∀>∃)

More explicitly, the Memory Buffer definition prevents any element from being remerged out of the originating phase (*non-local retrieval) or else copied/discharged in the memory buffer of a superordinate phase (*upward inheritance):
Then the clause-boundedness of QR follows from the computational sequencing of phases: the matrix subject QP1 is stored in the Q-buffer (Q1) of the containing phase P1 (the variable insertion in the argument position is ignored in the derivation). The embedded subject QP2 is stored in the Q-buffer of P2 (Q2), but it cannot “get into” the Q-buffer of the previously computed P1. As a result, QP2 will only have scope over the embedded P2, whereas QP1 will have scope over both P1 and P2. Thus, the phase boundaries determined by this top-down model, though not corresponding to complete subtrees, derive the clause-boundedness of QR, subsuming one instance of the otherwise mysterious "right roof constraint".

5.5. C-command and Pronominal Binding in a Top-Down Left-Right Grammar

As for QR, we can use the memory buffer device and the phase-based inheritance mechanism (no upward inheritance, no non-local retrieval) to implement any specific kind of relevant C-command relation. As we saw in (18)-(19) binding provides an important set of tests to detect the attachment point of an extraposed element.

Following Bianchi 2010 (phase-based extension of Schlenker’s 2005 proposal) we can implement A-binding using a Referential Buffer 25 that is a third type of memory buffer used to store the denotation of the referents (i.e. proper names, demonstrative pronouns and definite descriptions) that are processed during the derivation from left to right. The special properties of such buffer are:

i. Elements within the memory buffer do not need to be discharged in the end of the processing for the sentence to be grammatical;

ii. We admit some degree of freedom in retrieving the referents in the memory buffer (no rigid LIFO structure is assumed 26);

Below the schematic example on how referents are retrieved then evaluated:

---

25 We need to introduce some simplification here to keep the discussion focused on a minimal, relevant, set of facts. For this reason I will ignore how these referents are indexed and how the sequence of referents is used to evaluate the truth conditions of the sentence (see Schlenker 2005 and Bianchi 2010 for discussion).

26 This is because locality does not seem to play a role in binding as strong as in movement (but see Grosz et al. 1995).
When the referential expression \((\text{John}, \text{P}_2)\) is processed, its referent is stored (step 2) in a phase-local \(R\)\((eferential)\)-buffer \((\text{R}_1)\) which is different from the \(M\)(ove)-buffer and from the \(Q\)(uantifier)-buffer since:

a. it does not need to discharge/remerge the elements it contains at the end of the derivation;

b. both nested and selected phases inherit the R-buffer of the containing open phase (step 4);

The bound pronoun retrieves the referent from within the R-buffer (step 5) and it is evaluated before the phase is closed. Schlenker (2005) and Bianchi (2010) propose that whenever in the evaluation sequence (which roughly corresponds to our memory buffer) an element which is already present is reintroduced, we get a violation of the Principle of Non-Redundancy (i.e. Condition C violation).

5.6. Extraposition from a Left-Right, Top-Down perspective

Let us first distinguish EXT of arguments from EXT of adjuncts: in the case of EXT of arguments (“I gave him a painting yesterday of John” schematically depicted later in (65)), the “heavy” (i.e. containing select features) DP (“a painting of John”) competes with a right hand adjunct (“yesterday”, which is projected rightward because of a select feature in a specific functional position, as explained in §5.3, (58b)) for minimizing nesting (59). The four available solutions are:

i. Keep everything “in situ” (“yesterday we saw a painting of John”);

ii. Both shift the whole (heavy) adjunct and the whole (heavy) constituent (“we saw yesterday a painting of John”)

iii. Shift the adjunct and just phase projection of the heavy NP (“we saw a painting yesterday of John”)

iv. Just shift the adjunct and leave the heavy constituent is situ (“we saw a painting of John yesterday”)

We are not interested now in ranking or excluding (for independent reasons) some of the proposed solutions27. What is crucial, according to the discussion in §5.3, is that, in order to get the EXT version in iii., we need some movement trigger to keep Phase Projection local to the verbal head (thus guaranteeing a sort of Complement-Principle à la Guéron and May, 1984). Following Fox and Nissenbaum (1999), I propose that the relevant trigger is the quantificational status of the DP (nominal phase) that, once QR-ed (rightward, as discussed in §5.4, without any stipulation) is remerged in a peripheral position after the nuclear scope has been computed. There, it can satisfy the selection requirements of the phase head, first applying (the delayed) Phase-

---

27 Here, my intent is just to show how different derivations can be produced in the current framework and which one should be preferred following general principles (e.g. minimize nesting). A precise account of optionality is beyond the scope of this paper.
Projection, then Merging the relevant argument. Notice that the selection requirements of the nominal phase head is unsatisfied in situ and this prevents the relevant phase from being closed (the last selection requirement is not “phase projected”), and processed as discussed in §5.5. The content of the R-buffer is then available up to the point of the derivation in which the argument “of John” is sequentially merged. This produces a violation of Non-Redundancy as soon as the referent “John”, already present, since coreferent with “him”\(^{28}\), is inserted in the local R-buffer yielding a condition C effect.

\[(65)\]

The minimal difference with respect to EXT of adjuncts resides on the fact that the nominal phase “a painting” is complete/interpreted when QR takes place; then when “from John’s collection” is late-attached, once the NP has been remerged after QR, the evaluation sequence (i.e. the memory buffer) is no more available.

\[(66)\]

As in the original Guéron and May (1984) and Fox and Nissenbaum (1999) proposal, EXT is (“clause”)/phase-bounded since:

i. QR cannot scope out of the superordinate phase (Chesi and Bianchi 2010);

ii. An adjunct clause, in order to be a nested phase, needs to be attached to the first open (and compatible) superordinate phase.

For the very same reason we predict that, by default, embedded PPs cannot host EXT targeting a superordinate phase (the Q-buffer does not permit neither upward inheritance, nor non-local retrieval):

\(^{28}\) The evaluation of *him* requires that some individual be already present in the referential memory buffer once the pronoun is processed the first time. This is possible because the R-buffer is populated by individuals that are in the relevant conversational common-ground (Bianchi 2010).
5.7. Heavy NP-Shift
As briefly suggested in the paragraph before, Minimize Nesting (59) would predict also HNPS. Since selectional requirements have to be satisfied, by definition, in a local configuration with respect to the phase head, we simply predict “clause”-boundedness without any further assumption:

(68) a. Who did you [give the books written by the venerable Prof. Plum] to?

Notice that when the shifted phase is selected, it becomes the “last selected phase”, that is, we can predict movement “from” this constituent ((69a) vs. movement from the internal one (69b)):

(69) a. Who did you [give the books written by the venerable Prof. Plum] to?

b.
Since we can correctly predict that shifted selected constituents behave as last selected (sequential) phases, the “ne” extraction from focalized and heavy NPs is straightforwardly captured ((29).b vs. (30)).

5.8. Remaining issues

In the remaining pages I would like to sketch the tentative solution for two other issues that I did not have space to discuss in depth here: the ban on EXT from the canonical subject position in Italian and the availability of EXT from embedded constituent in Dutch and, sometimes, in English too.

We noticed in §2.2 that EXT from the subject position is not cross-linguistically uniform: English (10a), for instance, allows for such operation but Italian does not (12). As we saw in §3.2, Italian allows for HNPS of subject over the direct object and this is possible because of the availability of pro in such language (this makes the criterial subject position “recoverable” under (59)). As noticed by Calabrese (1982), Belletti and Shlonsky (1995), this shifting operation results in a focus on the shifted subject. Assuming that EXT is also linked to focalization (Huck and Na 1990, among others), we could predict that the status of the Italian canonical subject position (topicalized and not focalized 29) makes it incompatible with the status of “focalized constituent” that is necessary in order to allow EXT of an inner constituent 30. This is indeed possible under focalized conditions that seem to be the standard conditions for the subject in languages like English:

(70) a. Italian type languages (subject: +Topic)

As noticed by Culicover and Rochemont (1990), EXT from topocalized element is not allowed in English either:

29 In this respect, I assume, following Cardinaletti (2003), that the “canonical” subject position is SubjP: a functional projection below CP (FinP) and above TP where the “subject of predication” is licensed. This licensing is a “criterion” in the sense of Rizzi (2006).

30 Notice that HNPS of a subject is not blocked in Italian; this means that the pro strategy is potentially available to satisfy a criterial subject/topic position.
(71) a. John said he would meet a man at the party who was from Philadelphia, and meet a man at the party who was from Philadelphia he did.

b. *John said he would meet a man at the party who was from Philadelphia, and meet a man at the party he did who was from Philadelphia.

Another open issue is the availability of Extraposition from embedded constituents in languages like Dutch (17) and, sometimes, also in English as discussed in Strunk and Snider (2013). I am inclined to believe that such possibility is much more restricted than what has been reported in literature and it is essentially possible only from the last selected constituent (the last VP-shell):

Despite the fact that these data deserve more scrutiny, this option could be coherent with the nature of the last selected (sequential) argument since the last complement has some degree of independence with respect to the selecting phase. Then I would prefer not to weaken the Right Roof Constraint or the (generalized) Subjacency idea (as I reframed it in Top-Down terms), but rather concentrate the analysis on this very restricted set of subcases.

6. Discussion

In this paper, I propose that rightward movements, like Extraposition and Heavy NP-Shift, can be successfully accommodated in a Left-Right, Top-Down derivation (where every movement is to the right) without reducing their empirical peculiarities, which are clause boundedness, insensitivity of Extraposition to adjunct/argument asymmetry and the definiteness constraint. All of these properties can be explained using the notion of Phase Projection (a phase-based version of theta role assignment) and a new conception of memory-buffered long distance dependencies. Within this radically different perspective we are still able to characterize as finely as necessary such non feature-driven movements and to capture how they affect scope relations (e.g. selectively bleeding condition C effects, Fox and Nissenbaum 1999, following the derivational top-down implementation of binding principles proposed by Schlenker 2005, Bianchi 2010). With respect to Extraposition, I have argued that the quantificational status of the host is important for a relevant subset of phenomena (this can hardly be captured in a purely phonological way, e.g. Göbbel 2007),

31 “Müller's counter-examples [...] always involve extraposition along a chain of complements. Nonlocal complement extraposition from adjuncts still appears to be degraded, thus contrasting quite strongly with adjunct extraposition” (Crysmann 2013).
Do the “right” move

Cristiano Chesi

moreover, the directionality of QR does not need to be stipulated (Bianchi and Chesi 2010, vs. Fox and Nissenbaum 1999).

In this system, every long distance dependency (e.g. A'-movements, QR etc.) is regulated by specific triggers (e.g. unselected argumental features inserted in a functional position, the unavailability of a relevant domain to be computed, e.g. nuclear scope in the case of QR) using memory buffers, which are locally connected to the notion of phase and regulated by a simple discharge mechanism (only the last selected, sequential phase can inherit the content of a previous memory buffer). In this way we can capture the productive distinction between nested and recursive constituents (the rightmost selected complement) that allows for successive cyclic movements, with no need of look-ahead features to trigger intermediate steps. On the other hand, by relating heaviness to the presence of select features on a nested phase, we can predict both shifting and Extrapoosition as a result of the tendency to reduce nesting (computational complexity). This leads to marginalize “heavy” phases by casting the required phase projections to the end (i.e. to the right) of the phase. Crucial properties that characterize Extrapoosition and Heavy NP-Shift simply follow from the notion of inheritance of the memory buffers: they are phase-bounded (i.e. “right-roof constrained”) because phase-projection always takes place within the superordinate phase, and the memory buffers can neither be discharged upwards, nor on a non-selected phase.

It is important to stress, in this volume, that these empirical generalizations required a derivational assumption that contrasts with the Minimalist one, which is widely assumed to be “bottom-up”. In fact, assembling phrase structures Top-Down results in a phrase structure growth that is “from left to right”, since expectations must be fulfilled only after they are introduced in the derivation. This produces a precise derivational sequence that is compatible with the sequence in which the items are parsed and generated in the structure. Also crucial is the subdivision of the derivation in phases that are computational chunks driven by local selection and feature lexicalization. It is the geometry and sequence of phase expansions that permit to explain restrictions like the right-roof constraint that must be otherwise simply stipulated. Since the empirical generalizations here discussed seem to be correct and many apparently unrelated properties have now an integrated explanation, I think that this approach might be promisingly on the right track.

References


Why Left-to-Right Grammar?:
Evidence from Korean

Jieun Kiaer
University of Oxford
jieun.kiaer@orinst.ox.ac.uk

This paper aims to explain Korean native speakers’ strong preference for incremental, left-to-right structure building based on the following core phenomena: (i) left-right asymmetry; (ii) striking preference for early association; (iii) dislocation and context-sensitivity. I claim that those phenomena reflect the procedural aspects of linguistic competence, which are hard to capture/explain in a static grammar formalism. In this paper, based on these observations, I argue for the necessity of a grammar formalism which assumes left-to-right incrementality as the core property of syntactic architecture. Though the nature of discussion is theory-neutral, to formalise procedural competence, I will adopt Dynamic Syntax (DS: Kempson et al. (2001), Cann et al. (2005)) and provide an account based on DS. By adopting a left-to-right framework, I believe that we can bridge some unnecessary gap between COMPETENCE and PERFORMANCE, meeting the ultimate goal of any linguistic theory, that is, to achieve the appropriate explanatory adequacy. In this paper, I argue that structure building, both in comprehension and production, is driven by the native speaker’s need to optimise their structure, which is expressed as a strong tendency to build any meaningful, communicative proposition as quickly as possible with a minimised structure-building effort (cf. Hawkins (2004)).

1. Empirical challenge in the grammar

1.1. Beyond Descartes’ problem
The mystery of human language lies in the way humans construct their linguistic message in a systematic way across languages.¹ Even very small children can freely produce sentences they have never heard. This is known as Descartes’ problem. However, such freedom is not unrestricted. One of the core abilities of native speakers of a language is the ability to tell whether a given sequence of words is well-formed or ill-formed in their mother tongue. This provided strong evidence to generative grammarians that the logic of

¹ This paper is financially supported by British Academy Small Research Grant SG-49436.
structure, namely, syntax, constitutes the core part of human language. In this paper, I want to bring up another aspect of natural language syntax. It is that native speakers do not equally produce all the logically possible well-formed strings in real language use. In addition, they do not understand with equal efficiency all the logically possible well-formed strings. Instead, they show a strong preference, expressed as efficiency or frequency, for certain well-formed strings both in understanding and in their own production. The same tendency is also found in typological variation as well as in language change. Neither diachronic nor synchronic variation displays all the logically possible alternatives, but mostly a certain very limited set of variations.

This paper aims to answer why it is so and what it tells us about the architecture of mental grammar. Ever since Chomsky (1965), generative grammarians have believed the existence of core, abstract, universal and innate properties 2 of human languages used by an ideal speaker which could well answer the Descartes’ problem. I will call it core grammar in this paper for simplicity of discussion. This paper shares the same belief on the existence of the core grammar, yet questions the exact nature of such grammar. I claim that a strict and radical gap between innate knowledge of a language, known as competence, and behavioral variation, known as performance is the result of looking at the grammar from a static point of view, ignoring its procedural, time-sensitive and dynamic nature. I will show that such an unnecessary gap can be bridged and that we can provide a more sensible/elegant syntactic account by adopting a grammar with a left-to-right growth mechanism.

In particular, based on the three phenomena we will discuss, I will show dynamic aspects of the core grammar, which reflect procedural competence and are expressed as left-to-right incremental growth of a linguistic structure. I argue that left-to-right syntactic architecture is essential in explaining procedural competence, as it can naturally capture/explain incremental growth of a linguistic structure. Non-linear, static grammar formalisms are hard to pin down in those phenomena, since they are yielded as consequences of incremental growth of a structure, as we shall see.

Particular evidence is to be given from Korean, known as a strict verb-final and (relatively) free word-order language. Nevertheless, most of the issues and challenges that we will discuss are not just applicable to some very limited number of languages, but also can be extended to verb-final languages in general, which comprise about 40% of world languages according to the World Atlas of Language Structures (WALS) Online. Indeed, I believe that the properties that we will discuss are manifestations of natural language syntax. In this paper, I will focus on the following aspects and questions to examine the properties of the core grammar.

(1) (a) Freedom at Left and Restriction at Right

---

2 Similarly, Chomsky described I-language with the following terms: *the ‘pure’ manifestation of the language faculty that would arise in the mind/brain of a child placed in an idealized homogeneous speech community of one of the varieties of what we loosely and probably incoherently call ‘English’, etc.* (Chomsky 1987:31), recited from Wasow (2002:142).
Why are expressions at the left periphery interpreted more freely than those at the right periphery?
(b) Early Syntactic Association
Why do native speakers STRONGLY prefer early syntactic association and put the arguments in the same local domain?
(c) Dislocation Mystery and Context Sensitivity
Though early association is strongly preferred, why do native speakers sometimes use dislocation (even across a clause boundary)?

In the following, in Section 2, I will discuss each phenomenon and the result of some empirical investigations. In Section 3, I will introduce Dynamic Syntax (Kempson et al. 2001, Cann et al. 2005, Kempson and Cann in this volume) briefly and provide a DS analysis. In section 4, I will conclude.

2. Core phenomena

2.1. Freedom at left and restriction at right
In Korean, expressions at the left periphery can be interpreted more flexibly. For instance, a sentence-initial dative NP Komtoli-hanthey ‘to a bear’ in (2) can be interpreted in the three possible structures hosted by a verb malhaysseyo ‘said’ (= 2a), mantwule-cwukessta ‘make-give’ (= 2b) or yaksokhayssta ‘promised’ (= 2c).

(2) THREE POSSIBLE INTERPRETATIONS AT THE LEFT PERIPHERY
Komtoli-hanthey tokki-nun taramjwui-ka mass-iss-nun cake-lul
bear-dat rabbit-top squirrel-nom taste-exist-adn cake-acc
mantwule-cwukessta-ko yaksokhayssta-ko malhayss-eyo.
make-will.give-comp promised-comp said-decl
(a) ‘A rabbit said to a bear that a squirrel promised that he will make and give him a delicious cake.’
(b) ‘A rabbit said to somebody that a squirrel promised to a bear that he will make and give him a delicious cake.’
(c) ‘A rabbit said to somebody that a squirrel promised that he will make and give a bear a delicious cake.’

However, interpreting expressions becomes gradually more restricted. So, in (3), komtoli-hanthey can have just two possible interpretations. Yet, as it gets close to the verb, only one interpretation becomes available as in (4)-(5):

(3) TWO POSSIBLE INTERPRETATIONS
Tokki-nun komtoli-hanthey taramjwui-ka mass-iss-nun cake-lul
bear-top rabbit-dat squirrel-nom taste-exist-adn cake-acc
mantwule-cwukessta-ko yaksokhayssta-ko malhayss-eyo.
make-will.give-comp promised-comp said-decl
(a) ‘A rabbit said to a bear that a squirrel promised that he will make and give him a delicious cake.’
(b) ‘A rabbit said to somebody that a squirrel promised to a bear that he will make and give him a delicious cake.’
(c) ‘A rabbit said to somebody that a squirrel promised that he will make and give a bear a delicious cake.’

(4) **ONLY ONE POSSIBLE INTERPRETATION JUST BEFORE A MATRIX VERB**
Tokki-nun taramjwui-ka mass-iss-nun cake-lul mantwule-cwukessta-ko rabbit-top squirrel-nom taste-exist-adn cake-acc make-will.give-comp yaksokehayssta-ko komtoli-hanthey malhayss-eyo. promised-comp bear-dat said-decl
(a) ‘A rabbit said to a bear that a squirrel promised that he will make and give him a delicious cake.’
(b) ‘A rabbit said to somebody that a squirrel promised that he will make and give him a delicious cake.’
(c) ‘A rabbit said to somebody that a squirrel promised that he will make and give a bear a delicious cake.’

(5) **ONLY ONE POSSIBLE INTERPRETATION AT THE POST-VERBAL, RIGHT PERIPHERY**
Tokki-nun taramjwui-ka mass-iss-nun cake-lul mantwule-cwukessta-ko rabbit-top squirrel-nom taste-exist-adn cake-acc make-will.give-comp yaksokehayssta-ko malhayss-eyo komtoli-hanthey. promised-comp said-decl bear-dat
(a) ‘A rabbit said to a bear that a squirrel promised that he will make and give him a delicious cake.’
(b) ‘A rabbit said to somebody that a squirrel promised that he will make and give him a delicious cake.’
(c) ‘A rabbit said to somebody that a squirrel promised that he will make and give a bear a delicious cake.’

Though Korean is regarded as a strictly verb-final language, it is very easy to find right-peripheral expressions. According to the picture-description test that we will discuss in 2.2, when the 7-9 year old children want the dative NP to be interpreted within an embedded clause, 87% of the time, they placed dative NP either just in front of a verb or after a verb. See Figure 3.

Yet, in such cases, they are STRONGLY preferred to be interpreted in its closest local structure. Therefore, (6a) sounds very odd and ungrammatical, unlike (6b). Kiaer (in press) shows that in the grammaticality judgement test, 70% of the time, native speakers (n=33) found sentences as in (6b), where the dative NPs can be incorporated to the closest leftward structure, very natural. In these cases, the dative NPs cannot be incorporated to the closest leftward structure, since the host verb is a transitive verb.

(6) a. Ditransitive verb and transitive verb sequence PLUS dative NP

<table>
<thead>
<tr>
<th>H-top</th>
<th>J-nom</th>
<th>sakwa-lul</th>
<th>cwuessta-ko</th>
<th>saynggakhayssseyo</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hena-nun</td>
<td>Jina-ka</td>
<td>gave-comp</td>
<td>thought</td>
<td></td>
</tr>
</tbody>
</table>

3 n in this paper means the number of participants.
Mina-hanthey.
M-dat
Intended reading: ‘Hena thought that Jina gave an apple to Mina.’
b. Transitive verb and ditransitive verb sequence PLUS dative NP
Hena-nun Jina-ka sakwa-lul mekessta-ko malhayssayo Miina-hanthey.
H-top J-nom apple-acc ate-comp said M-dat
Intended reading: ‘Hena said to Mina that Jina ate an apple.’

Post-verbal expression and the verb often form an intonational unit with the preceding verb as Park (2003) showed. Park (2003) discussed that between the verb and its post-verbal argument, the Intonational Phrase (IP) boundary tone\(^4\) may be shared or copied. See the pitch tracks from Park in Figure 1 for the example (7).

(7) Etise mandu-ni? ku-kapang
where make-Q the-bag
‘Where did they make it, the bag?’

Figure 1: Pitch Track from Park (2003)

The circled line is drawn by me and it refers to an intonational unit between the verb and its post-verbal argument. As we can see, the boundary marking (%) doesn’t occur after the verb but after the post-verbal NP.

Why do expressions at the left or sentence-initial position have more freedom in terms of choosing its host structure, compared to the expressions at the right or sentence-final position? Since the freedom of interpretation decreases GRADUALLY from left-to-right as a structure unfolds, it is difficult to answer the question within a binary feature-based framework as we shall discuss more in 2.4.2. The fact that structural freedom decreases from left-to-right, regardless of the location of a verb, clearly shows us that a syntactic

\(^4\) In this paper, IP refers to the Intonational Phrase.
structure is being built from left-to-right, rather than waiting for a verb and remaining un-built.

Left-right asymmetry of structure building has been a puzzle in the grammar, ever since it has been first discussed by Ross (1967) as a Right Roof Constraint. However, it occurs to me that indeed the phenomenon as such is not the reflex of specific aspects of Korean syntax, but instead it is a manifestation of the general/universal aspect of linguistic structure building. This component of syntactic competence is impossible to capture if we disregard the left-to-right growth property of a syntactic structure.

2.2. Early syntactic association
In the last section, we have observed that expressions occurring at the left periphery have more freedom than those at the right periphery. However, the empirical data we will discuss in this section show that native speakers STRONGLY prefer to interpret or understand the sentence-initial, left-peripheral expression within the FIRST-available, CLOSEST LEFTWARD structure in the left-to-right, time-linear understanding – instead in ANY arbitrary structure. Consider (2) again.

(2) THREE POSSIBLE INTERPRETATIONS AT THE LEFT PERIPHERY
make-will.give-comp promised-comp said-decl
(a) ‘A rabbit said to a bear that a squirrel promised that he will make and give him a delicious cake.’
(b) ‘A rabbit said to somebody that a squirrel promised to a bear that he will make and give him a delicious cake.’
(c) ‘A rabbit said to somebody that a squirrel promised that he will make and give a bear a delicious cake.’

Sentence-initial dative NP CAN be interpreted in all three structures hosted by mantwule-cwukessta ‘made-and-give’, yaksokhayssta ‘promised’ and malhayss-eyo ‘said’. Yet, native speakers of Korean STRONGLY prefer (8a) reading over all possible readings. According to Kiaer (in press), when native speakers of Korean (n=33) were asked a question such as (8b) after hearing (8a), 96% of the time, they answered in such a way which implies the sentence-initial dative NP being interpreted in the matrix clause together with the following topic marked NP.

(8) ONE DOMINANT INTERPRETATION AT THE LEFT

made.gave-comp said
(i) ‘A rabbit said to a bear that a squirrel made a delicious cake for him (= rabbit).
(strongly preferred: 96% of the time)
(ii) ‘A rabbit said that a squirrel made a delicious cake for a bear.’
‘Who did the squirrel make the cake for?’
If the answer is a bear: (i) reading is chosen.
Yet, if the answer is a rabbit: (ii) reading is chosen.

In the following, at first I will provide a series of evidence which shows native Korean speakers’ strong preference for early syntactic association both in comprehension and production. Then, I will provide the procedural competence these phenomena reflect.

**Comprehension-and-Production Test** The aim of the test was to investigate whether native speakers of Korean **UNDERSTAND** and also **SPEAK** incrementally, by observing their interpretation of a dative NP *nwukwu-hanthey* ‘to whom’. Two groups have participated in the test. One group was high school students (aged 17-18, n=118) and the other group was primary school students (aged 7-9, n=20). Both groups were asked to look at the picture and answer the question as in (9).

![Illustration](image)

(9) Condition A: Dative NP Subject (NOM) Subject Object Verb Verb
(i) ‘To whom did the zebra say that the giraffe was giving an apple to the elephant?’
(II) ‘To whom did the giraffe give an apple, according to what the zebra said to the rabbit?’ (Expected answer: an elephant)

Condition B: Dative NP Subject (TOP) Subject Object Verb Verb
nwukwu-hanthey ellukmal-un khirin-i sakwa-lul cwun-ta-ko
In principle, sentence-initial dative NPs can be interpreted in the same or different clause with the immediately following NP. Yet, the result of this comprehension test shows in both groups native speakers strongly prefer to interpret the dative NP in the same clause, signalled by the immediately following NP, though in principle they can be interpreted freely in other clauses. (As for high school students, such tendency was observed 87% for condition A and 86% for condition B and as for primary school children, such tendency was observed 97% for condition A and 93% for condition B of the time.) No significant effect was observed between Condition A and B.

In this comprehension test, the clause, where both the dative NP and the following nominative or topic-marked NP were interpreted, was the matrix clause. Yet, one cannot generalise that the first-available structure for the sentence-initial dative NP is always a matrix clause. Consider (10). Even if the first two lexical items are the same with the Condition A sentence of (9), the sentence-initial dative NP and the following nominative NP are most likely to be interpreted in an embedded clause this time. That is, (10a) reading is much preferred to (10b), unless the prosodic break intervenes between the sentence-initial dative NP and the following nominative NP.

(10) First-available structure is the embedded clause
Nwukwu-hanthey khirin-i sakwa-lul cwun-ta-ko ellukmal-i who-dat giraffe-nom apple-acc give-decl-comp zebra-nom malhayss-e?
said-Q
(a) ‘To whom did the zebra say that the giraffe was giving an apple to the elephant?’ (Expected answer: a rabbit)
(b) ‘To whom did the giraffe give an apple, according to what the zebra said to the rabbit?’ (Expected answer: an elephant)

After the comprehension test, the same subjects participated in the production test. As for high school students, they had to write a question, inducing the answer written already in their questionnaire. The answer was either the recipient animal or the animal listening to the story. As for primary school children, being paired in two, they were instructed to ask a question to the other student, inducing the answer circled in their pictures. The answer could also be either the recipient animal in the bubble or the animal listening to the story. Primary school children’s question/answer practices were recorded for further prosodic analysis.
Both groups’ production patterns show some interesting results. Let’s first look at the result from high school students. (11) shows preferred structural sequences and preferred syntactic association patterns in those sequences.

(11) Early Association at Left Periphery

<table>
<thead>
<tr>
<th>Left-peripheral Sequences</th>
<th>Occurrence</th>
<th>Preferred Association</th>
</tr>
</thead>
<tbody>
<tr>
<td>(a) Nom-Dat-Nom</td>
<td>424</td>
<td>{Nom-Dat}-Nom</td>
</tr>
<tr>
<td>(b) Dat-Top-Nom</td>
<td>134</td>
<td>{Dat-Top}-Nom</td>
</tr>
<tr>
<td>(c) Top-Dat-Nom</td>
<td>1486 (37% of the total)</td>
<td>{Top-Dat}-Nom</td>
</tr>
</tbody>
</table>

Although ambiguous readings are plausible (i.e., immediate or non-immediate reading) in NP sequences given (11a-c), native speakers UNANIMOUSLY preferred to associate the dative NP to the ALREADY available, LEFTward structure, unfolded by proposition-unfolding particles such as nominative or topic-marked NPs, not by a verb. Besides, about 45% of the time (1807 occurrences), the participants placed dative NP before either an embedded or a matrix verb. In these cases, the construal of dative NPs is not ambiguous: dative NPs are to be interpreted in the on-going propositional structure, to be completed by the upcoming verb.

Let’s also look at the data we gathered from 7-9 year old children. As for the primary school students’ data, when the answer animal was that of listening to the story, inducing the matrix clause reading of the wh-dative NP, the wh-dative NP occurred mostly in the left-periphery (83% of the time) as in Figure 2. (S1 means the first lexical item in the sentence.)

![Figure 2: Left-peripheral Dative NPs: Matrix Clause Reading](image)

Yet, when the answer animal was the recipient animal inside a bubble, inducing the embedded clause reading of the wh-dative NP, the wh-dative NP occurred most likely just before a verb or after a verb (87% of the time) as in Figure 3. In those cases, the participants were most likely to pronounce a verb and a post-verbal lexical item in the same intonational unit, as Park (2003) observed.
What is very striking here is that in both groups even if left-peripheral dative NPs can be interpreted freely in any level of embedding, native speakers strongly prefer to use left-peripheral dative NPs when they intended the matrix clause interpretation, yet right-peripheral dative NPs when they intended the embedded clause interpretation. In this way, it seems that native speakers avoid any possible ambiguity. I claim that such disambiguation in production is the consequence of resource-sensitive structure building along with early syntactic association at the left periphery. In other words, these production patterns show that native speakers do NOT arbitrarily build a syntactic structure but do it in the most efficient way. Here some fundamental question could arise. In the tenet of generative grammar, an ideal speaker is implicitly assumed to have the capacity to understand and speak all the logically possible forms with similar efficiency. Under this assumption, it is hard to explain why certain forms are more efficient and therefore more frequent than other possible forms. Such efficiency/frequency effects are often ignored or regarded as only “accidental”. Yet, it is counter-intuitive to see the above results being merely coincidental. Rather, I believe the data we discussed manifest the linguistic competence of Korean native speakers and beyond.

At the left periphery, where structural ambiguity could arise, native speakers of Korean preferred to resolve the dative NP in the FIRST AVAILABLE propositional structure. This structure is not necessarily a matrix clause or a particular level of embedded clause as we have seen. It is the first available, CLOSEST LEFTWARD structure in the left-to-right, time-linear understanding.

Yet, the first available, CLOSEST LEFTWARD structure can be defined in left-to-right architecture only. So, theorising syntactic competence, expressed through early syntactic association is not easy in non-linear, static grammar formalisms. The syntactic competence we observe in this section is that native speakers make syntactic association incrementally just by having a sequence of particles, without the full knowledge of the upcoming structure. To capture/explain early syntactic association and the notion of first available structure, a grammar formalism should assume left-to-right incremental growth of a structure in its backbone. In the case of Korean, I propose that the constructive use of particles and some routinised structure-building processes enable such incremental
structure building. Particularly, I argue that nominative particles (-ka/-i) and topic particles (-un/-nun) are those particles which unfold a proposition. I will return to this in 3.3.

Furthermore, I argue that early syntactic association, indeed evidenced by synchronic, diachronic, and typological syntactic variations (See Kiaer (in press)), results from the following goal in syntactic structure building. The insight for (12) is from Hawkins (2004). I will return to this in 3.1.

(12) Goal of Linguistic Structure Building:
(For an efficient communication), native speakers aim to optimise their syntactic structure building (both production and comprehension) by achieving the meaningful, communicative proposition as quickly as possible with minimised structure-building effort.

2.3. Dislocation mystery and context sensitivity
We have seen in the last section native speakers’ strong tendency for early syntactic association. Here, a puzzling question arises. If it is so, then why is it the case that long-distance dependency or any form of dislocation ever exists in all languages. In particular, if we assume (12), long-distance dependency should NOT exist, because it causes syntactic association being delayed. Nevertheless, even if those forms look inefficient, they ARE used and sometimes PREFERRED in certain cases. In this section, we will discuss why it is the case and what it tells us about the grammar.

Most grammar formalisms concentrated on how to explain long-distance dependency formation, neglecting the question on why such seemingly inefficient syntactic phenomena widely exist across languages. To answer this question, we need to extend the target of natural language syntax to be beyond a single-sentence level.

Sentences with dislocation are not preferred in general. As we have seen in the last section, NONE of the time, native speakers produced long-distance dependency in their production. Yet, when an appropriate context is given where the current structure NEEDS to be linked/associated to receive an interpretation, context-sensitive or anaphoric expressions are often fronted and dislocated from their local structure, even across a clause boundary. Consider (13) and (14).

(13) Context:
  a. Do you know the tall person who comes to the library every day?
  b. To that tall guy, Jina said that she has once lent her chemistry notebook.
  c. Jina said that she has once lent her chemistry notebook to that tall guy.

(14) Context:
  a. Tosekwane mayil o-nun khi-ku-n saram al-ci?
     library-at everyday come-and height-tall-adn person know-Q
     ‘Do you know the tall person who comes to the library every week?’
Why Left-to-Right Grammar?  Jieun Kiaer

(13c) like structures are much more frequent than the structures like (13b). Likewise, (14c) like structures are much more frequent than the structures like (14b). Nevertheless, when a context such as (13a) or (14a) is given, (13b) and (14b) become even more natural than (13c) and (14c). (14b) is an example with long-distance dependency, where an embedded clause subject as well as the complementiser -ko is dropped and the matrix verb is simplified. Examples like (14b) are regarded difficult to understand in the generative grammar (Saito 1992), yet indeed are very natural and easily observed in spoken Korean.

In fact, context-sensitive or anaphoric items tend to occur at the left periphery – sometimes, even across a clause boundary. According to the 10-million words Sejong corpus search, the topic-marked anaphoric pronoun ku-kess-un or its shortened form ku-ken ‘the-thing-top’ have been found 7,734 times. Strikingly, in all times ku-kess-un or ku-ken (the-thing-top) hardly appeared in the sentence-medial, final or right-peripheral positions, though in principle they CAN occur at ANY place. Most times, they occurred between the two sentences – at the end of one sentence and the beginning of the other sentence, where their accessible context is close. Among the 7,734 cases, I have extracted only two examples where the other argument precedes ku-kess-un or ku-ken ‘the-thing-top’ at the beginning of a new sentence. This result directly reflects context-sensitivity of natural language syntax. In a strict sense, no syntactic account can be given to why ku-kess-un or ku-ken ‘the-thing-top’ strongly prefers to occur at the left-periphery.

In a single-sentence level, one may attempt to provide a feature such as +front to those lexical items. Yet, note that though they strongly prefer to be located at the left-peripheral position, the sentences with right-peripheral or sentence-medial ku-kess-un or ku-ken ‘the-thing-top’ do also yield a GRAMMATICAL sentence. Just as in left-to-right asymmetry phenomena, given that context-sensitivity is the general aspect of natural language syntax, rather than an attribute from a set of lexical items, if we try to provide a feature-based account, we will then end up with numerous unnecessary features.

I argue that the motivation behind the leftward dislocation, known as long-distance dependency, is to maximise the use of previous context in current syntactic structure building in order to increase efficiency of communication. Leftward dislocation provides a way of building a CROSS-sentential syntactic structure in an incremental manner. By placing context-sensitive, anaphoric expressions ahead, hearers can quickly resolve the semantic value of those expressions and efficiently enrich their on-going structure through the preceding context.

5 I did not regard this position as being right-peripheral, since those expressions were not incorporated into the previous sentence, but into the upcoming sentences.
Consider another pair of examples given in (15)-(16). In both examples, (b) sentences are odd, unlike (a) sentences, due to gender mismatch. In Koran, some family relational words are gender sensitive. Enni ‘sister(F)’ should be used when it refers to a girl’s sister. On the other hand, nwuna ‘sister(M)’ should be used if it refers to a boy’s sister. Both (b) examples are bad because Jina is a typical girl’s name in Korean. Hence, instead of nwuna ‘sister(M)’, enni ‘sister(F)’ should have been used.

(15) Gender mismatch in a simple sentence (b)
      sister(F)-dat J-top note-acc lent-decl
      ‘Jina lent a note to her sister(F).’
   b. ???Nwuna-hanthey Jina-nun note-lul pilrieuwuess-e.
      sister(M)-dat J-top note-acc lent-decl
      ‘??Jina lent a note to his sister.’

(16) Gender mismatch in a complex sentence (b)
      sister(F)-dat J-top note-acc lent-said
      ‘Jina said that she lent a note to her sister(F).’
   b. ???Nwuna-hanthey Jina-nun note-lul pilrieuwuess-tay.
      sister(M)-dat J-top note-acc lent-said
      ‘??Jina said that she lent a note to his sister.’

(15) is a simple sentence and (16) is a complex sentence with long-distance dependency. Yet, regardless of the level of embedding or the existence of long-distance dependency, the grammaticality judgement of both pairs of sentences seems to be made when the gender mismatch is observed at the beginning of a sentence. This is clearly far ahead of the time when information such as the type of a sentence (i.e., declarative, question, relative clause, etc.) or the number of embedding becomes available. It will be only at the very end of a sentence when native speakers recognise any sort of long-distance dependency. Consider (17). In (17b), even though native speakers understand the sentence-initial dative NP Enni-hanthey ‘sister(F)-dat’ as the argument of the embedded clause, they cannot resist interpreting Enni-hanthey ‘sister(F)-dat’ and the following NP Jina together in the same clause unfolded by a topic-marked NP Jina-nun. This is also why (17c), which should be grammatical is in fact ungrammatical.

(17) a. Grammatical Sentence without Dislocation
   J(F)-top K(M)-nom sister(M)-dat note-acc lent-said
   ‘Jina said that Kiho lent a note to her sister.’

   b. Should be Ungrammatical, yet Grammatical
   sister(F)-dat J(F)-top K(M)-nom note-acc lent-said
   ‘Jina said that Kiho lent a note to her (Jina’s) sister.’

   c. Should be Grammatical, yet Ungrammatical
The gender mismatch phenomena we see above shows that even grammaticality judgement occurs incrementally from left-to-right. It also shows that a syntactic structure is built incrementally regardless of the type of a structure or the number of embedding. Also, no matter whether a whole sentence is composed of a simple clause or a complex clause, native speakers tend to front the given lexical item to help the listeners to link the two sentences. In this way, dislocation/long-distance dependency can also increase the efficiency of real-time understanding and can be understood as the extension of left-to-right growth of a syntactic structure at the cross-sentential level. This seems to be the basis of dislocation or what is known as long-distance dependency. Therefore, it is hard to grasp the driving force of dislocation or long-distance dependency in a single sentence level.

2.4. Explaining core phenomena: the limit and challenge
In the last section, we have observed the three core phenomena. They can be summed up as the consequences of incremental, left-to-right structure building both in a sentential and cross-sentential level, aiming for optimised syntactic structure building.

In the following, I will discuss some approaches made in grammar formalisms to capture/explain incrementality. In explaining incrementality in the core grammar, as a way to decrease any uncertainty in structure-building, Aoshima et al. (2004) have proposed un-forced revision and McConvile (2001) adopted the notion of ruthless parser. Yet, as we shall discuss, those accounts are basically limited in explaining what is driven by left-to-right, incremental growth of a structure.

2.4.1. Approaches in derivational grammars
Phillips Challenge Phillips (1996, 2003) argued that even defining a constituent may be difficult without taking linear order into consideration. The same string of words can be diagnosed as a constituent by one constituent test, but not by another constituent test. (18) shows a case of constituency conflict. (18) is from Phillips (2003). When co-ordination is taken as a test for constituency, as in (18a), *Gromit a biscuit* passes the co-ordination test and thus, is regarded as a constituent. Yet, as in (18b), taking permutability as a criterion, *Gromit a biscuit* cannot pass the movement test and thus, is not regarded as a constituent.

(18) Constituency Conflict

<table>
<thead>
<tr>
<th>[Gromit a biscuit] is a constituent by co-ordination test</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. Wallace gave [Gromit a biscuit] and [Shawn some cheese] for breakfast.</td>
</tr>
<tr>
<td>[Gromit a biscuit] is NOT a constituent by movement test</td>
</tr>
<tr>
<td>b. *[VP Gomit a biscuit] Wallace gave VP for breakfast.</td>
</tr>
</tbody>
</table>

Based on constituency shift and constituency conflict phenomena, Phillips argued for the Incrementality Hypothesis as in (19).

(19) Sentence structures are built incrementally from left-to-right.
As shown in Figure 4, Phillips assumed that syntactic relation must respect constituency at the point in the derivation when the relation is established. Yet, once this relation is licensed, constituency (between A and B) may change subsequently (i.e., be revised); and this he argued was the basis for such conflict.

![Figure 4: Incremental Constituency via Build-and-Revise Process](image)

Given that defining constituency is a crucial matter in syntax, it is striking that such a core notion is difficult to sustain without considering left-to-right growth of a structure. Yet, explaining incrementality of a verb-final language is not so straightforward as we shall see.

**Problems in Aoshima et al. (2004)** In this section, I will particularly discuss problems that may arise in the generative grammar, in the attempt to explain left-to-right incremental growth of structure building.

Generative grammar-formalisms are clearly limited in their ability to explain incrementality displayed in verb-final languages like Korean/Japanese, because the logic of the formalism is not sensitive to the left-to-right linear order.

Aoshima and her colleagues (Aoshima et al. 2003, 2004) have argued that in Japanese the sentence-initial *wh*-dative NP undergoes *un-forced revision* during the on-line structure building, to be interpreted within the ‘most’ embedded clause. Consider their examples:

(20) **Scrambled, Declarative Complementiser: Delay at -to**

<table>
<thead>
<tr>
<th>Dono-seito-ni</th>
<th>tannin-wa</th>
<th>koocyoo-ga</th>
<th>hon-o</th>
</tr>
</thead>
<tbody>
<tr>
<td>which-student-dat</td>
<td>class-teacher-top</td>
<td>principal-nom</td>
<td>book-acc</td>
</tr>
<tr>
<td>yonda-to</td>
<td>tosyositu-de</td>
<td>sisyo-ni</td>
<td>imasita-ka?</td>
</tr>
<tr>
<td>read-decl</td>
<td>library-at</td>
<td>librarian-dat</td>
<td>said-Q</td>
</tr>
</tbody>
</table>

‘Which student did the class teacher tell the librarian at the library that the principal read a book for?’

(21) **In-situ, Declarative Complementiser: Delay at -to**

<table>
<thead>
<tr>
<th>Tannin-wa</th>
<th>koocyoo-ga</th>
<th>dono-seito-ni</th>
<th>hon-o</th>
</tr>
</thead>
<tbody>
<tr>
<td>class-teacher-top</td>
<td>principal-nom</td>
<td>which-student-dat</td>
<td>book-acc</td>
</tr>
<tr>
<td>yonda-to</td>
<td>tosyositu-de</td>
<td>sisyo-ni</td>
<td>imasita-ka?</td>
</tr>
<tr>
<td>read-decl</td>
<td>library-at</td>
<td>librarian-dat</td>
<td>said-Q</td>
</tr>
</tbody>
</table>

‘Which student did the class teacher tell the librarian at the library that the principal read a book for?’
Scrambled, Question Particle: No Delay at -ka
Dono-seito-ni tannin-wa kooeyoo-ga hon-o
which-student-dat class-teacher-top principal-nom book-acc
yonda-ka tosyositu-de sisyo-ni imasita.
read-Q library-at librarian-dat said
‘The class teacher told the librarian at the library which student the principal read a book for.’

In-situ, Question Particle: No Delay at -ka
Tannin-wa kooeyoo-ga dono-seito-ni hon-o
class-teacher-top principal-nom which-student-dat book-acc
yonda-ka tosyositu-de sisyo-ni imasita.
read-Q library-at librarian-dat said
‘The class teacher told the librarian at the library which student the principal read a book for.’

Aoshima et al. found that both when the wh-phrase was in situ as in (21) and (23) or when they are scrambled long-distance as in (20) and in (22), Japanese speakers were surprised to encounter a declarative complementiser -to in the embedded clause. Based on this result, they argued that Japanese speakers at first posit a gap in the matrix clause but revise it into the most embedded clause. According to Aoshima et al., such shift is “unforced”. Nevertheless, they still assumed that the motivation of such revision is feature-checking. In other words, Aoshima and her colleagues have argued that the slowdown in the embedded clause with a declarative marker is caused because of the failure in feature-checking, in particular, in failure to check the wh-feature. Aoshima et al. (2005) argued that this is why referential NPs in a sentence initial position do not undergo unforced revision, in contrast to wh NPs. However, given that the strength of wh-feature in Japanese/Korean is very weak (Grewendorf 2001), it seems strange that native speakers perceive them so radically different. It is also hard to see why native speakers can and do make distinctions between two types of NPs for the future structure building at such an early stage of a structure building.

Aoshima et al. (2004) also argued that the parser prefers to satisfy requirements brought from the initial dative NP as soon as possible, hence, within the most embedded clause, since in Japanese/Korean the most embedded verb comes earliest and the matrix verb comes last. However, without getting the full picture of a structure, it is impossible to know how many embedded structures are used in the given sentence. Hence, we cannot see from which source structure such revision should take place until we get the full picture of the sentence. Aoshima et al.’s account is what generative grammars such as GB/MP can at best offer, yet is inadequate to capture/explain the essence of incremental, left-to-right structure building.

---

6 Aoshima et al. (2004:42) noted, “By unforced reanalysis, we mean a revision that is not licensed by any incompatibility of the initial analysis of the parse with subsequent material.”
**Right Periphery Puzzle**  Ever since Ross (1967), locality restriction at the right periphery have been widely discussed across languages. Lasnik and Saito (1992) proposed the Proper Binding Condition to capture the ungrammaticality of (24c). According to this analysis, (24c) is derived from two-step movements: the first step is (24a) to (24b) and the second step is (24b) to (24c).

(24) a. \[TP Taroo-ga [Hanko-ga Sooru-ni iru-to] omotteiru\] (koto)
   \[Taroo-nom [Hanako-nom Seoul-in be-that] think\] fact
   ‘Taroo thinks that Hanako lives in Seoul.’

b. \[TP Sooru-ni [Taroo-ga [Hanko-ga \(t_i\) iru-to] omotteiru]\] (koto)
   \[Sooru-nii [Taroo-nom [Hanako-nom be-that] think\] fact
   ‘In Seoul, Taroo thinks that Hanako lives \(t_i\).’

c. *\[TP Hanko-ga \(t_j\) iru-to]j [Sooru-ni [Taroo-ga \(t_j\) omotteiru]]\] (koto)
   \[[Hanako-nom be-that] [Sooru-nii [Taroo-nom think\]]] fact
   ‘[That Hanako is \(t_j\) in Seoul Taroo thinks \(t_j\).’

In (24a), the PP *Sooru-ni ‘in Seoul’* is first scrambled from the embedded clause to the matrix initial position as in (24b). And then, the embedded CP \[TP Hanko-ga \(t_i\) iru-to\] is scrambled to the position preceding the PP *Sooru-ni* in (24c). The resulting sentence (24c) is totally ungrammatical. According to Saito, the ungrammaticality is caused by the unbound trace \(t_i\) in (24c). The problem, however, is that the level of S-structure no longer plays any role in movement frameworks. Indeed any restriction that relies on S-structure becomes completely inexpressible in minimalist explanations, even given Saito’s own account of radically reconstructed sentences. This is because after radical reconstruction, no trace should be left at LF. The Proper Binding Condition was a S-structure condition, but in the Minimalist Program, S-structure doesn’t exist any more. Saito (2003) explained the ungrammaticality of (24c) by reformulating the Proper Binding Condition as a constraint on the application of Merge. Merge combines two linguistic objects to form a constituent. The two objects to be combined by Merge must be a ‘complete’ constituent. A constituent which has only a part of a chain, e.g. a trace but not its antecedent, does not qualify as a constituent. In (24c), \[TP Hanko-ga \(t_j\) iru-to\] is disqualified as it is not a complete constituent. Hence, this sequence cannot be subject to Merge (See Saito (2003) for a detailed discussion). However, given the pro-drop property of Korean and Japanese, it is not clear how to decide whether a constituent is complete or not.\(^7\)

The Proper Binding Condition debate shows a limit to non-linear grammar, capturing the property derived from left-to-right growth of a structure. Since the nature of phenomena is purely derived from left-to-right growth of a structure, grammars which are not sensitive to linear-order face the limit of providing explanatorily adequate accounts on these phenomena (See Chesi, in this volume).

**2.4.2. Approaches in lexical grammars**

\(^7\) Moreover, unlike Japanese, in the case of Korean, expression could also occur after the complementiser. In picture-description task (2.2), 47 times of the time, primary school participants placed \(wh\)-dative NPs after the complementiser *-ko.*
Explaining left-to-right, incremental growth of a structure in lexicalist grammars such as Combinatory Categorial Grammar (CCG), Head-driven Phrase Structure Grammar (HPSG) or Lexical Functional Grammar (LFG) is not so straightforward. See Kiaer (2007) for a detailed discussion.

First of all, in these grammars, it is assumed that the verb provides the most important combinatory information in structure building. So, within these grammars, it is impossible to assign any provisional structure in an incremental manner before the verb is reached. Consequently, it is hard to capture incremental structure building of verb-final languages in these grammars.

Among these grammars, Combinatory Categorial Grammar (CCG) particularly assumes incremental, step-by-step growth of a structure. Nevertheless, as McConville (2001) pointed out, it is hard to capture the heart of incremental, left-to-right growth in CCG, since, though it assumes step-by-step, derivational procedures, they are not sensitive to linear order. To capture incremenotality in CCG, McConville assumed a ruthless parser to reduce any structural uncertainty. Yet this approach may not work in verb-final languages like Korean/Japanese, since some crucial structure-building information only occurs at the very end. Though native speakers build a syntactic structure incrementally, there is NO way for them (= hearers) to make the right prediction on the type of a sentence or the number of embedding for the current sentence at the beginning or left-peripheral position. Yet, just as in generative grammars, to yield a well-formed derivation in CCG, human parsers/native speakers must know the whole structure at hand at the start of structure building.

I believe that the elegant syntactic theory should be able to reflect the PARALLEL judgements that we have observed between a simple sentence without dislocation as in (15) and a complex sentence with dislocation as in (16), repeated below.

(15) Gender mismatch in a simple sentence (b)
      sister(F)-dat J-top note-acc lent-decl
      ‘Jina lent a note to her sister(F).’
   b. ???Nwuna-hanthey Jina-nun note-lul pilriecwuess-e.
      sister(M)-dat J-top note-acc lent-decl
      ‘??Jina lent a note to his sister.’

(16) Gender mismatch in a complex sentence (b)
      sister(F)-dat J-top note-acc lent-said
      ‘Jina said that she lent a note to her sister(F).’
      sister(M)-dat J-top note-acc lent-said
      ‘??Jina said that she lent a note to his sister.’
Yet, in a non-linear grammar formalism, it is not easy to capture this. Baldridge (2002) among others showed how multi-modal CCG can indeed provide a step-wise derivation for sentences with long-distance dependency. But, he proposed two separate analyses for a short and long-distance scrambling. Particularly, as for a long-distance scrambling, he proposed some complicated type-shifting for the cross-sentential, dislocated lexical item. Yet, given that native speakers of Korean seem to understand both sentences in the same, incremental way, an analysis to posit different syntactic composition for the same sequence of lexical items seems to lack explanatory adequacy.

Karamanis (2001) modified CCG and specified the argument’s syntactic position in the lexicon to capture linear-order sensitivity in Greek. For instance, the value $+FRO$ denotes that there is a fronted argument involved. Yet, such approach is again, I believe, far from an explanatorily adequate account. Notice that if we adopt Karamanis’ approach, the size of lexicon will be uncontrollable in relatively free word-order languages. In addition, the left-ness/right-ness is a relative and gradual notion, which cannot be captured using a binary notation such as $+FRO$. Feature-based approaches therefore will inevitably face the same problem with Karamanis (2001).

**2.4.3. Towards grammar for procedural competence**

What the three core phenomena reflect is the time-sensitive and also resource-sensitive nature of linguistic competence, which I will call the PROCEDURAL COMPETENCE in this paper. The procedural competence is particularly observed by (i) left-to-right and (ii) optimized structure building (See (12)). As we have seen in 2.4.2, however, capturing/explaining both properties in grammar formalisms is not easy. Firstly, it is because of the long-held assumption between competence and performance: For instance, distributional variation between grammatical forms has been regarded as nothing to do with the core grammar, no matter how significant the distributional asymmetry is, whereas any asymmetry between grammatical and ungrammatical forms have been considered relevant in shaping the core grammar. Hence, in a strictly orthodox view of generative grammar, the phenomena we have observed in this paper will not even be regarded as what syntactic theory should explain. Secondly, as we have observed, non-linear, static grammars are inadequate to explain the left-to-right growth property of natural language syntax.

In order to capture procedural competence, what we need is a grammar which can capture the left-to-right growth of a structure. In other words, we need a grammar which has the left-to-right growth mechanism as the backbone of its syntactic architecture. Also, we need to have a grammar which can explain why only a very limited syntactic choice becomes dominant in real use.

Though the nature of discussion is theory-neutral, to formalise procedural competence, I will adopt Dynamic Syntax (DS: Kempson et al. (2001), Cann et al. (2005)) and provide accounts for the core phenomena in the next section.
3. Explaining core phenomena in dynamic syntax

3.1. Dynamic syntax: Left-to-right challenge
Dynamic Syntax (DS) assumes left-to-right directional derivation. Such assumption makes it possible to thread left-to-right incremental growth of a structure into the heart of the core grammar. Following Sperber and Wilson (1995), DS takes as the starting point in any structure-building, the goal to establish some PROPOSITIONAL formula as interpretation. In this paper, I revise the goal as proposed in (12), which incorporates resource-sensitive nature of natural language syntax.

(12) Goal of Linguistic Structure Building:
(For an efficient communication), native speakers aim to optimise their syntactic structure building (both production and comprehension) by achieving the meaningful, communicative proposition as quickly as possible with minimised structure-building effort.

That is, I argue that the goal of linguistic structure building is not just to obtain a proposition but to do so in the most optimised way. Later in this section, I will show how such optimisation process is encoded in the grammar.

The essence of DS architecture lies in the use of UNDERSPECIFICATION. It assumes that a structure is underspecified at the beginning but is updated/more specified following the linear order as more information is incorporated. This approach is close to Phillips’ (1996) claim (See (19).) Yet, rather than assuming that a structure is build-and-revised from left-to-right as in Phillips, DS assumes that a structure is underspecified-and-updated/specified along the time line. By adopting structural underspecification and its subsequent update we can explain left-to-right asymmetry as observed in (2)-(4). Intuitively, at the beginning of a structure building, a provisional structure will remain largely underspecified. Yet, as it progresses, it will be more and more updated/specified. Freedom in structure building will therefore decrease as a structure grows.

In DS, three ways of structure building are assumed: (i) local/immediate update (via an operation of local *adjunction); (ii) non-local/non-immediate update (via an operation of *adjunction)

Simply speaking, DS assumes that structure can be further developed in its on-going local structural template or in the non-local template or in an arbitrary template in principle. (See Kempson et al. (2001) and Cann et al. (2005) for a formal detail.) The availability of these three options can well-capture flexible structure building at the left periphery.

However, as the structure-building progresses, not all options will become available. Particularly, at the right periphery, after the verb is reached, since a structural skeleton is

8 In DS, the Kleene star (*) operator is used to characterise an underspecified tree relation. The essence of this relation is similar to the concept of functional uncertainty adopted in LFG (Kaplan and Zaenen 1989). The modality <↑> is an underspecified modal relation, which can be extended into any functor (1) or argument (0) relation.
already completed or in the process of being sealed off a local/immediate update option becomes available. This can well-capture why expressions at the right periphery have limited interpretations unlike those at the left periphery.

In addition, though in principle all three operations are available in its original set-up as in Kempson et al. (2001) and Cann et al. (2005), I argue that native speakers ROUTINISE the operations in a way that the most restricted option (i.e., immediate growth) is most preferred, unless other factors such as context or prosody are intervened as given in (25).

(25) Routinised Rules:
Immediate Growth >> Non-immediate Growth >> Arbitrary Growth
‘Every child ate his/her apple.’

Before going on further, I will show some basic logic of DS.

3.2. Basic logic in DS
A basis for modelling the core architectural properties of DS\(^9\) is LOFT (Logic Of Finite Tree; Blackburn and Meyer-Viol (1994)), as illustrated in Figure 5. In the LOFT, there are there core functors. First, \(Tn (= \text{TREE NODE})\) functor. This indicates the address of a tree; Second, \(Fo (= \text{FORMULA})\) functor. This is regarding the content of a node; Third, \(Ty (= \text{TYPE})\) functor. This has all the combinatory information of a node. Consider Figure 5. In this tree, the pointer, , indicates a node under development. \(?Ty(t)\) indicates the overall goal of establishing a proposition of \(Ty(t)\) with the sub-goal of a predicate, \(?Ty(e \rightarrow t)\). The concept of requirement, \(?X\) for any \(X\), is the central concept of the framework: all requirements, whether of content, structure, or type, have to be met by the end of a sequence of parse actions. For example, the requirement for an update of content (\(?\exists xFo(x)\)) in an argument node, needs to be met either by update from within the structure-building process or from context. Consider (26) and the structure projected by Bomi’s answer (underlined) as in Figure 5.

\(^9\) For more detailed formal account, see Kempson and Cann (this volume) and Cann et al. (2005).
Why Left-to-Right Grammar?  Jieun Kiaer

(26) Hemi: Semi-ka  Mina-lul  manna-ss-ni?
Hemi: Semi-nom  Mina-acc  meet-past-Q
‘Has Semi met Mina?’

Bomi: Manna-ss-e.
Bomi: meet-past-decl
‘(Yes, Semi) met (Mina).’

(1) Ty(t), Tn(a)

(2) ?∃xFo(x)
Fo(U) = Semi :
(updated from the context),
<↑0>Tn(a)

(3) ?Ty(e→t)

(4) ?∃xFo(x)
Fo(V) = Mina :
(updated from the context)
<↑0><↑1>Tn(a)

(5) Fo(manna),
Ty(e→(e→t))

Figure 5: Propositional Structure Projected by manna (= ‘meet’)

At first, let’s see the top node, labelled as ?Ty(t), Tn(a). This is the node (1) in Figure 5. What ?Ty(t) indicates is that a proposition (?Ty(t)) is required (?). This is the overall goal of structure building. In fully pro-drop languages like Korean/Japanese, a verb on its own can project a propositional structure. From just the parsing of the verb, the argument structure can be projected and then subsequently, the content values of the arguments will be updated by the context. Tn(a) can be simply understood as an underspecified S node in the sense that its level of embedding whether it is a matrix or an embedded clause is underspecified.

Now, let’s move from the top-node to the next node along with an argument (0) relation. This is the node (2). The node under 0 relation in the LOFT is the argument node and the node under 1 relation in the LOFT is the functor node. The argument node of {?Ty(t), Tn(a)} is labelled as {?∃xFo(x), Fo(U), <↑0>Tn(a)}. This node is to be a subject node since it is the argument of a proposition. ?∃xFo(x) can be read that a formula (or content) is required for this node. In Fo(U), U is a place-holding meta-variable, whose content is to be updated by the context. <↑0>Tn(a) refers to the modal relation, which states that from the current node, if the parser goes up (↑) along with 0 (= argument) relation, there is a topnode (= Tn(a)), which is a proposition/sentence. The other daughter node under ?Ty(t), Tn(a) is the predicate (= VP) node. This is the node (3). This node is the functor argument of a proposition. ?Ty(e→t) indicates that a predicate is required. This node can be addressed as <↑1>Tn(a).

From this node, two sub-nodes are further developed. First, let’s move to the node along with the argument (0) relation from the ?Ty(e→t) node. That is the node (4). This node is to be an object node. This node is labelled as {?∃xFo(x), Fo(V), <↑0><↑1>Tn(a)}. Fo(V)
is another place-holding meta-variable, projected by the verb, whose value also needs to be identified in the context. \(<↑_0><↑_1>Tn(a)\) is the tree node address, which states that if the parser goes up along the argument (= 0) node relation and then the functor node (= 1) relation, there is a top-node (= Tn(a)). Another argument node branched from \(?Ty(e→t)\) is the verb node. This is the node (5). shows the current state of parsing. This node can be addressed as \(<↑_1><↑_1>Tn(a)\).

Notice that the structure in Figure 5 is built/projected by a verb alone. Given that a language like Korean is a fully pro-drop language, what this structural template assumes is that a partial structure can be projected even in the absence of arguments as their content values can be retrieved from somewhere in the parsing process or from the context. In the case of above structure, the content values of a subject and an object are retrieved from the context question. Yet, as we shall see in the next section, contrary to the above example, where a provisional structure is unfolded by a predicate, a partial structure can be unfolded by case morphology even before a VERB is reached. This, we believe, in fact is the source of incremental structure building in head-final languages like Korean.

### 3.3. Incremental Syntax via Constructive Nominal Particles

In this section, I will show how DS can explain incremental structure building, even in the absence of a predicate. As we have seen in Figure 5, a propositional array CAN be unfolded with a verb alone. Nevertheless, it is not just a verb which can project a structural template. If it is so, incremental structural growth in verb-final languages will altogether remain as a serious problem: Will the native speakers of more than 40% of world languages understand/speak backwards? Kiaer (2007) discussed the constructive role of case particles in Korean, following the insight from Nordlinger (1998). Following these previous works, in this paper I argue that native speakers can incrementally build a partial structure by the use of case particles along with routinised options of structure building that we will discuss in the next section. (27) show what syntactic actions are encoded by the nominal particles. The topic markers -nun/-un has a more complicated structural anticipation. Yet, I will leave this out in this paper.

<table>
<thead>
<tr>
<th>Particles</th>
<th>Syntactic Anticipations</th>
</tr>
</thead>
</table>
| -ka/-i (NOM) | a. Unfold a propositional structure  
b. Be the subject of that structure |
| -nun/-un (TOP) | a. Unfold a propositional structure that is to be the root structure  
b. Be the subject of that structure |
| -lul/-ul (ACC) | a. Attach to the closest leftward structure  
b. Be the direct object of that clause |
| -hanthey (DAT) | a. Attach to the closest leftward structure  
b. Be the indirect object of that clause |
Suppose that a sequence of case-marked NPs such as (NOM plus ACC or NON plus DAT) are parsed. The partial structures that we can obtain will be as follows:

<table>
<thead>
<tr>
<th>NOM + ACC</th>
<th>NOM + DAT</th>
<th>ACC + DAT</th>
</tr>
</thead>
<tbody>
<tr>
<td><img src="image.png" alt="Diagram" /></td>
<td><img src="image.png" alt="Diagram" /></td>
<td><img src="image.png" alt="Diagram" /></td>
</tr>
</tbody>
</table>

Figure 6: Partial Structures Projected via Case Particles

Through the constructive use of case particles, a partial structure will be unfolded as above. This partial structure will then be merged with the complete proposition projected by a verb.

**Verbal Cluster: Completion of Structure Building**  Though Korean is known to be a verb-final language, expressions occur so frequently after the verb. As Park (2003) initially discussed, in such a case, the verb and the post-verbal expression are likely to form an intonational unit. This has been observed in the recording of primary school children’s question/answer pairs we discussed in 2.2. Given this empirical observation, I argue that a verb cluster at the right periphery contains not only the complete propositional template, but also prosodic information such as tone which can confirm the completion of the structure-building. For instance, a clear IP boundary tone (See Jun (2000)) could indicate the ending of a question, yet when such tone is missing at the verb cluster, the hearer would hold on the completion of the on-going structure building. At this juncture, a post-verbal expression with the closing-off IP boundary tone could be inserted at the not-yet-completed structure and seal-off the current structure. (28) shows that what syntactic anticipations are encoded in the grammar by the verbal cluster.

(28)

<table>
<thead>
<tr>
<th>Verbal Cluster</th>
<th>Syntactic Anticipations</th>
</tr>
</thead>
</table>
| First part of a verbal cluster: root PLUS tense | a. Unfold a complete proposition  
b. Update the previously unfolded partial proposition |
| Second part of a verbal cluster: sentence-ending particle PLUS closing-off IP boundary tone  
Sentence-ending particle PLUS hesitant tone | Confirm the completion  
Do not yet complete the sentence |
However, even if the role of case particles is crucial in unfolding a sentence, they can be so easily dropped. In particular, native speakers of Korean most likely drop case markers when the context makes the role of its argument obvious and clear. However, when the role of an argument becomes not obvious, the use of a direct object particle becomes natural. It is noticeable that in the spoken data collection of 7-9 year old children that we discussed in 2.2, not a single time did children drop case particles when they were asking a question to the other child, whereas they invariably dropped the case particle when they were answering the questions.

**Sample Structure Building**  
Consider (29). (29a) is a simple sentence (without cross-clausal dislocation) and (29b) is a complex sentence (with cross-clausal dislocation). % refers to the IP boundary.

(29)  
a. ???Nwuna-hantehy Jina-nun note-lul pilricwuess-e.  
sister(M)-dat Jina(F)-top note-acc lent-decl  
‘Jina lent a note to his siter.’  
b. ???Nwuna-hantehy % Jina-nun note-lul pilricwuessta-ko saynggakhay.  
sister(M)-dat Jina(F)-top note-acc lent-comp thought  
‘Jina thought she lent a note to his siter.’

Let’s first think of how a structure is built for (29a). At first, the sequence of three pre-verbal NPs will yield the following partial structure given in (30). This partial structure then will be merged with the structural template projected by a verb later. If an inappropriate verb occurs, such structure building will fail (e.g., transitive or intransitive verbs). In principle, each structural relation can have three possible choices (i.e., immediate update, non-immediate update, arbitrary update). Yet, according to the routinisation as given in (25), immediate update was chosen in all structural relations. See Kiaer (2007) for a detailed step-wise derivation process.

(30) Partial structure built via *Nwuna-hantehy Jina-nun note-lul*
Notice that this partial structure is semantically awkward, because of the gender mismatch. That is, Fo(Nwuna) requires +MALE, yet Fo(Jina) is –MALE. This shows that not structure building but also anaphoric resolution or any semantic update occurs in an incremental way.

Now, let’s look at (29b). The difference between the partial structures build by (29a) and (29b) lies only in the fact that the structural relation for the dative NP is unfolded by the non-immediate relation (via *adjunction). The sentence-initial dative NP is to be interpreted in the embedded clause. Such long-distance dependency is most likely to be indicated by prosodic break or the given context. This makes the native speaker UNAMBIGUOUSLY choose to build the NP in a non-immediate clause over an immediate cause, which is a default option. The partial structure built before a verb will hence be different as below. Nevertheless, even if the dative NP is not syntactically resolved in this partial structure, due to the immediate anaphoric update for Fo(Nwuna) through Fo(Jina), the sequence still remains awkward. Just as in a simple clause structure-building, after the verb cluster is parsed, the whole structure-building will be ready for completion. Unless we assume incremental anaphoric resolution along with incremental structure building, we cannot explain semantic awkwardness of both (29a) and (29b).

4. Conclusion: Towards grammar for procedural competence

In this paper, I discussed Korean native speakers’ strong preference towards incremental, left-to-right structure building and proposed the procedural competence which can bridge the unnecessary gap between competence and performance. Based on some empirical observations, this paper argues for the necessity of a grammar formalism which assumes left-to-right growth as the core property of syntactic architecture. In addition, I proposed that a core grammar should be able to capture/explain structural optimisation and the resource-sensitive nature of natural language syntax. In particular, I claim that structure building is driven by the native speaker’s need to optimise their structure building by achieving the goal of linguistic building – to achieve a meaningful communicative proposition as quickly as possible with minimised structure-building effort.

5. References

Chesi, C. (this volume) Do the ‘right’ move.


Syntax and Symmetry

Michael Brody
UCL, LI-HAS

1. Mirror Theory (MT) methodology
A plane, like the surface of an infinitely large lake has infinitely many symmetries. Rotation with any center and any angle, reflection with any mirror line, any translation results in the same plane. The set of symmetries of concentric circles, like circular waves created by a stone thrown in the lake, is a proper subset of the symmetries of the quiescent infinite lake-surface. Here rotation is symmetric only if its center is the stone's entry point, reflection only with a mirror line that traverses this entry point. No translation or other rotation or reflection is symmetric. Many symmetries of the plane are broken. We instinctively look for an explanation not of the symmetries but of symmetry breaking when we explain the concentric waves with reference to a dropped stone or a jumping fish. (The example comes from Stuart 1995.)

In general it is the departures from symmetry rather than the symmetries that are in need of explanation. Hence (validly) eliminating asymmetry from the theory is ipso facto making the theory more explanatory. This was the thinking behind earlier work in MT (Brody 1997, 2000) which went some way towards reintroducing symmetry into morpho-syntactic structures. However, if such an approach to syntax is taken seriously to its logical conclusion, it is in apparent opposition to all minimalist type asymmetrical structure building (including now also MT), which invariably involves a(nti)symmetrical concepts like labelled merge, spec-head, head-complement, probe-goal and c-command.

2. MT content
MT reduced the bare phrase structure (Chomsky 1994) tree in (1), where I indicated both the word (X0)-internal and the word-external structure, to the simpler (2), which involves only a morphological (eat-, -s) and a (symmetric) syntactic (John, -s) spec-head relation. The syntactic head-complement relation between -s and eat- was taken to be a (symmetrically) directionally mirrored morphological spec-head relation.

Given that word-internal spec-head relations were thus expressed in the form of syntactic complementation relation, the syntactic difference between heads and phrases (minimal and maximal projections, - whether absolutely or contextually defined) was eliminatable: instead of the redundant (circled) projection line in (3) a single representant of the morpheme -s serving both as head and as phrase suffices in (2). MT appeared to eliminate the need for phrasal projection. (Note that bare phrase structure trees eliminated only the absolute distinction between a head and phrase, but not the need for the projection line.)
3. Problems with MT

Let us first observe that the spec-head relation (or spec-x relation to be slightly more neutral, since the head in MT corresponds to both the head and the phrase of the minimalist approach), the basic primitive relation of MT, while construable in various ways, is inherently necessarily not symmetric, — a crucial problem for our present more radical approach. Secondly let us note a dissimilarity: the morphological spec-x relation is a pure configurational relation of concatenation, while the syntactic spec-x relation appears to involve feature sharing. A more precise representation of (2) is (2'), where some feature \( f(John) \) of the syntactic spec \( (John) \) shows up on the head -s.

\[
\text{(2') } \quad \text{-s } f(John) \quad \text{-s } \text{eat-}
\]

This observation by itself is just that, a fact. It leads however to a problem with two interrelated aspects both for MT and for the approach that I am currently pursuing. Spec-x agreement introduces a redundancy, it causes a feature to be repeated on another node. With the elimination of the projection line MT had some success in eliminating such redundancies, so the reoccurrence of a similar redundancy elsewhere is unwelcome news. In fact the problem seems worse. Spec-x feature sharing does not appear to be a concept genuinely different from phrasal (feature-)projection: in both cases a (set of) feature(s) of a node shows up on another (immediately dominating) node. There is little evidence to show that standardly assumed differences between the two relations are more real than apparent. If so, then the elimination of projection in MT is at best partial: the feature duplication relation underlying this notion shows up elsewhere.

To repeat, from our current perspective we have at least three problems for MT:
(a) spec-of-x is itself an asymmetric notion
(b) spec-x agreement still introduces redundancy
(c) spec-x feature sharing does not appear to be a concept genuinely different from (feature-)projection, and if so, then the elimination of projection is incomplete
4. Starke (2001)

Michal Starke proposes the "doubly filled nothing" restriction (DFN), according to this an XP projection contains either a spec and a complement, or a head and a complement but never both a spec and a head. His proposal can perhaps be phrased more perspicuously by saying that the (X'-projectional) feature(s) of a head H only show up on the node immediately dominating H, they do not percolate further to the node dominating the spec. So according to his hypothesis, instead of the standard tree in (3), in accordance with the DFN we would have the one in (4).

(3) 
```
  ● f(Spec), f(Head)  
    'Spec' ● f(Head)  
    'Head' Complement  
```

(4) 
```
  ● f(Spec)  
    'Spec' ● f(Head)  
    'Head' Complement  
```

The diagram in (4) throws into relief the identity of the spec-x feature sharing and that of phrasal feature projection, – a point made in the previous section. Starke considers the spec in (4) as just another head that projects and considers the notion of spec and spec-head relation thereby eliminated from the theory. We can indeed say spec's are heads since they project. But with equal justification we could interpret (4) by considering the head to be a spec since it is in spec-x agreement with the head. The choice appears to be no more than that of perspective and terminology. If we take the latter terminology then it is obvious that Starke's approach is not as different from MT as it might seem at first. Both approaches can be validly thought of as taking the spec-x relation as primitive. (The two theories interpret the x-complement relation differently. MT takes x-complement to be an alternative (morphological) type of spec-x relation while in Starke's framework the head-complement relation appears to be implicitly composed of two different primitives: a spec-x (projection) and an x-complement relation.)

Starke's theory adds also the DFN restriction, which also in MT terms entails that in a spec-x configuration x can only contain a subset of the features of the spec and no other features. So adopting the DFN in MT would result in expanding (2)/(2') as (5):

(5) 
```
  f(John)  
    John -s  
      eat-  
```
More importantly for our current discussion, Starke's approach does not fare better than standard MT with respect to the problems listed in (a),(b),(c) in section 3. above. I adopt here Starke's terminology of head-complement relations instead of the spec-x based wording to make this clearer. First, instead of the asymmetry of the spec-x relation, in Starke's framework we have the asymmetry of the head-complement relation. Secondly, just like spec-x agreement introduces redundant features into the representation, so does the X'-projection of features of the head in Starke's approach (the exact same phenomenon by a different name, if we are correct). And thirdly, if spec-x feature sharing is not distinct from (X'-feature-)projection, then Starke's elimination of the 'spec of' relation is incomplete: in the guise of projection/label it remains as a subpart of the head-complement relation. To repeat, corresponding exactly to the MT problems in (a), (b) and (c) in section 3, Starke's theory raises the following questions:

(a') head-complement an asymmetric concept
(b') projection (of features of the head) still introduces redundancy
(c') if spec-x feature sharing is the same as (X'-feature-)projection, then elimination of the 'spec of' relation is incomplete (in the guise of projection/label it remains as a subpart of the head-complement relation)

5. Symmetric sisterhood/Merge

It appears then, that our problems are caused by the mechanism that copies a set of feature(s) of a node to an immediately dominating node, sometimes called spec-head agreement, sometimes X'-theoretic head-feature-projection. Labelling is another common way to refer to the same concept. The logical and natural move is then to eliminate labeling, adopting the proposal of Collins (2002). Suppose we adopt symmetric sisterhood/Merge, the minimal assumption, and the obvious one given our general methodology. Clearly, an unordered set is a simpler, and therefore more preferable concept than an ordered one.

This makes all approaches which use the spec-head and/or the head-complement relation in syntactic structure building impossible, – almost all current approaches, including of course the standard minimalist system, Starke's approach and MT. (Note that the lack of a syntactic head-complement relation suggests strongly that neither s- nor c-selection is syntactic, these presumably belong to the semantic and morphological component, respectively.)

It is not a new assumption that the order of functional heads (complements) is interpretively constrained. If true, then the interpretive constraints that entail this ordering may make syntactic labelling (alias spec-x projection/agreement, alias X'-featural projection/agreement) redundant in the general case. This is particularly easy to see in Starke's framework where all non-terminal nodes immediately dominate a head and a complement. In a configuration like (6) for example, either node N1 or N2 projects/labels the immediately dominating node. Suppose N1 is or contains (in a way accessible for projection) head1, N2 head2, and N3 head3. Suppose further that the (interpretively determined) order of these heads is: head1 > head2 > head3. It follows, that in (5) head2 must be higher than head3, hence head2 and not head3 projects/labels the immediately dominating node. Thus there is no need to indicate by a label that head2 is head of (head2, head3) and head3 is complement of head2.
Given more standard assumptions about X'-theory, certain additional complications are introduced if the features of the head typically project also higher than the immediately dominating node and hence some nodes carry the features of both the spec and the head. Let us reasonably assume that the issues that arise here can be resolved, perhaps via the adoption of Starke's DFN and assume with Collins that there is no need to syntactically label in general.

But there are at least two major new problems that result from this step. First, symmetrical merge creates an apparently unsurmountable obstacle to linearization. All theories of linearization exploit one way or another the asymmetry introduced by the labeling mechanism of merge, either directly or indirectly by making reference to concepts like head, spec, complement or projection levels. Members of unordered sets apparently cannot be ordered in a linguistically relevant fashion. Secondly, given labelling/projection/spec–x agreement, all constraints that involve c-command can refer to the domination relation instead; as argued first in the context of MT, Brody 2000, – now see also Chomsky 2005. It seems obviously necessary to have a linearization algorithm for syntactic structures and it would clearly be desirable to retain the result that c-command reduces to simple domination.

6. Interpretive labeling – semantic order

Suppose that labelling for c-command is interpretive, and takes place in the semantic component. Then the same must be true more generally of c-command/domination constraints. Schlenker (2004) convincingly argues that principle C should be derived from a Gricean maxim of minimization whose effect is to eliminate what he analyses as pragmatically redundant restrictors, providing independent evidence that c-command asymmetries may indeed be interpretive in nature.

Schlenker proposes that "As a sentence is processed, top-down ..." "Each time a pronoun or an R-expression which denotes d is processed in a context c, its sister is evaluated with respect to c ∧ d, which is the context c to which d has been added. In other words processing an R-expression has the effect of making it 'super-salient' for the expressions that are contained within its sister. He argues that: "super-salient entities" (whether extra-sententially (linguistically or not) or intra-sententially provided by the above procedure) by the Gricean minimization principle "must be denoted using a pronoun, unless some special pragmatic effect is obtained by using a full description."

An important aspect of Schlenker's approach from our perspective is the fact that the interpretive top-down processing of the sentence results in a linearly ordered "sequence of evaluation", where if x precedes y in a sequence of evaluation then x is super-salient for y. But instead of adding an R-expression to a structure independent sequence of evaluation for elements contained in its sister node, we can add (the relevant features of) the R-expression to the node immediately dominating it and
Consider the node dominating the R-expression to be a member of the sequence. In other words, the sequence of evaluation for a category x is always an (ordered) sequence of interpretively annotated (labelled) nodes starting from the root of the tree and containing all nodes dominating x.

Note that there seems to be no need to restrict super-salience to binding theory: we may assume that other restrictions that in standard terms involved c-command make reference to this notion. In particular, we may assume that chains, which I have argued are interpretive constructs (e.g., Brody 1998), are similarly restricted: each chain-member must be super-salient for the next. C-command type asymmetries can continue to reduce to domination.

7. Interpretive labeling - linear order

Symmetric syntax is not compatible with mirror theoretical structures where a head dominates its complement. Nothing prevents however the assumption that interpretive labelling, which as we have seen appears to play a crucial semantic role, is also made use of in the morpho-phonological component. Let us make the natural assumption that labeling in the morpho-phonological component is by (morpho-)phonological features: the PF features of a bound morpheme head x are inserted in the node immediately dominating x. This effectively recreates mirror theoretical representations, where the morphemes of a word are adjacent elements; each (morpheme of a given word) being immediately dominated by (being the complement of) the linearly following one.

Such representations make a straightforward ordering algorithm possible, whose core principle is optimally simple:

\[ \text{if } x \text{ immediately dominates } y \text{ then } y \text{ precedes } x \]

We assume that where ID is a word-external relation (7) orders nodes of the tree, and where ID is a word-internal one (7) orders morphemes. Let us also additionally adopt (8) for expiciteness:

\[ \begin{align*}
\text{(8) a. } & \text{ID entails adjacency (of morphemes word-internally, of nodes/domains word-externally)} \\
\text{b. } & \text{all nodes must be ordered} \\
\text{c. } & \text{words are spelt out in their designated (highest?) node}
\end{align*} \]

(8a) can be thought of as a version of the no crossing branches requirement, explicit or implicit in all ordering theories. (8b) makes the order linear. (8c) establishes the link between morpho-phonological material and syntactic positions in a way that is apparently much simpler than standard head movement (see Brody 2000 for discussion).

Consider for example (9), a simplified core sentential structure. Given morpho-phonological interpretive labeling the PF features of the bound morpheme -s are copied to the immediately dominating node, resulting in the structure in (9'). As the PF features will be ignored in their source position by spellout (I indicate this by
parentheses in the diagram) the relation between (9) and (9') is in fact PF movement. I number the N nodes in (9') for ease of reference only.

Let us determine first the word-internal order of the morphemes -s and like-. We might proceed either by ignoring the morphologically and phonologically inert N nodes or by assuming that they form null parts of words. For presentational purposes I assume the latter, as this avoids some, apparently technical, difficulties. In (9') N₁ ID the morpheme -s, which ID N₂, which ID the morpheme -like, so by (7) the word-internal order is: like-, N₂, -s, N₁.

Next, we turn to the word-external, 'syntactic' order of words/positions. Here it is necessary to first determine the structural location of the word likes, whose morphemes occupy more than one position in the tree. The standard mechanism for establishing the relation between word-internal morphemes and word-external positions is head-movement supplemented by the highly arbitrary strength features. As indicated in (8c), following a MT proposal I tentatively assume that words are invariably taken to be located in the highest position their elements occupy, thus in (9') likes is spelt out in the position of N₁.

Using now (7) to determine word-external order, we see that John precedes N₁, since this node immediately dominates John. We have thus established that likes, spelt out in N₁, follows the subject. Furthermore, since all nodes must be ordered (8b), and the domain of any daughter of a given node N must be adjacent to N (8a), it follows that the other daughter of N₁, -s, must follow N₁. In fact, given the adjacency requirement (8a), everything dominated by -s will also necessarily follow N₁, and thus, by transitivity, will follow also John. As likes is spelt out in N₁, it precedes Mary, since the node Mary is dominated by -s; and -s and everything -s dominates follows N₁. Thus we have the correct word-external order: John precedes likes precedes Mary.

Consider the slightly more complex case in (10), (10') after labelling:
There is a word-internal relation between -s and have-, and between eat- and -en. The word-internal order, given the ID relations is have- precedes -s and eat- precedes -en. The former word is spelt out in the position of -s, the latter in that of -en. The ID relation between have- and -en is word-external, hence eaten spelt out in -en will precede the position of the have- node. But eaten in fact follows has. We can ensure this result if the have- node and its domain, hence also eaten, which is included in this domain, follows -s, the spell-out position of has. We can cause our principles entail that the have- node follows the -s node in (10') in exactly the same way in which they required the -s node follow N1 in (9') if the (-s) node in (10') is visible and the (-s) – -s relation counts as word-external.

We must not take however the relation between have- and (have-) to be word-external in (10') since that would result in have- immediately dominating two elements with which it would be in a word-external relation. This would create a contradiction: both elements would have to precede have- (7) and both would have to be adjacent to it (8a,b). Capitalizing on the difference between the -s – have- (word-internal) and the have- -en- (word-external) relation, we might ensure that (-s) – -s but not (have-) – have- is a word-external relation using the rather natural assumption in (11). According to this, elements that form no word-external relations (i.e. are word-internal morphemes) must be terminals of the tree.

(11) A non-terminal node N must ID some element with which it is in a word-external relation

Since the relation between -s and have- is word-internal, (11) entails that the (-s) – -s relation is word-external. Given (11) or some equivalent device, the invisible (-s) node in (10') will have to precede -s, with which it forms a word-external relation. Hence the -have node and its domain must follow -s, by (8ab). So eaten, dominated by -have will follow -s, the node where has is spelt out; and we have derived the order: has precedes eaten.
6. Bare checking, – the elimination of probe-goal asymmetry

The minimalist checking theory Chomsky (1994) is asymmetric: a c-commanding uninterpreted (and unvalued) probe searches for a c-commanded interpreted goal (with a value) to inherit its value from.

As is well known there are cases where the c-commanding probe appears to be interpreted raising an apparently significant problem for this theory. A typical example is (12) where it would seem that an interpreted +WH probe searches for a goal wh-phrase.

(12) I wonder who(+wh) C(+WH) Bill saw t
   +interpreted probe (?)

Pesetsky and Torrego (2004) point out, that Chomsky needs to posit "two distinct features in C: an uninterpretable, unvalued feature uWh with an EPP property (the feature that probes for a wh-goal); and a distinct, interpretable, valued feature iQ (the feature relevant to the interpretation of the clause). Correlations between clausal semantics and wh-type must be captured with mechanisms other than Agree" (p.7).

Their alternative proposal is to separate interpretability and valuation. In other words they assume that the +WH probe in (12) is interpretable but unvalued, and receives its value from an uninterpreted but valued wh-phrase goal.

We might wonder if the feature duplication of the earlier solution has not been simply exchanged for a different but perhaps equally unnecessary distinction between interpretability and valuation. Both duplications are made necessary by the assumption of asymmetry in the probe-goal configuration: that an uninterpreted (Chomsky) or unvalued (Pesetsky and Torrego) probe searches its c-command domain for a goal that provides interpretation or value for it. The suspicion that both duplications may be redundant is reinforced by the observation that it is an apparently unnecessary additional restriction that makes it necessary to invoke them.

Bare checking theory (Brody 1997b) simply requires all instances of features with a single semantic interpretation to be linked. Thus the interpretable +WH and the non-interpretable wh-feature of the wh-phrase must be (chain-)linked, there is no need to invoke an interpretation-independent notion of valuation. We do not need to talk about valuation, but if we wish to do so, we can allow the interpreted feature to value the non-interpreted one, – as expected on general grounds given the fact that valuation is semantically redundant. Without the ad hoc proviso that the structurally higher element must be non-interpretable or unvalued, we do not need to further duplicate our featural inventory invoking two sets of features for the analysis of such construction. See Brody 2000b for further discussion.

7. In a minimal symmetric syntax Move/Chain is undefinable

Returning to interpretive labelling, this copies a (set of) feature(s) to the immediate dominating node, where ID relation is understood in the standard way: x ID y iff x dominates y and there is no z such that x dominates z and z dominates y. In other words the ID relation (of interpretive labelling) implies a locality condition. The same locality condition is of course also implicit in the ID relation involved in the standard definition of Merge/sisterhood. Standard Merge creates a node that is in an ID
relation to the merged categories. If A and B are sisters then there is a node that is in an ID relation with both.

We might ask if the duplication of invoking ID/locality in both syntax and interpretation is indeed necessary. Suppose A and A’ are merged (sisters) and are (immediately) dominated by A'"; similarly B and B’ are (immediately) dominated by B". When we merge A" and B" creating say C, we take C to be the set whose members are (only) A" and B". Suppose however, that we understand Merge differently, so that merging the trees A" and B" will still result in C, but C is now defined as the set of all nodes it (non- reflexively) dominates and not directly by the ones it is in an ID relation with. In other words C dominates/contains A, A’, A"', B, B’, and B" and C is defined as the set whose members are exactly these elements.

An interesting consequence of this apparently innocuous simplification for the sake of eliminating a central redundancy is that the trees in (13) - (17) cannot be distinguished. They all represent the same single structure, drawn in different ways. Thus (13) for example contains two non-terminal nodes D, and C where C is defined as the node that dominates A and B while D is the node that dominates A, B and C.

(13) \[
D \\
| \\
C \\
A \quad B
\]
C: \{A, B\}
D: \{C, A, B\}

(14) \[
D' \\
A \quad C \\
A \quad B
\]
C: \{A, B\}
D’: \{C, A, B\} = D

(15) \[
D'' \\
B \quad C \\
A \quad B
\]
C: \{A, B\}
D’’: \{C, A, B\} = D’ = D

(16) \[
D' \\
C \\
A \quad B
\]
D’ = D

(17) \[
D'' \\
C \\
A \quad B
\]
D’’ = D’ = D

176
Given the standard set-theoretical assumption that the set \( \{ A, A \} \) is the same set as \( \{ A \} \), the tree in (14) contains exactly the same terminal and nonterminal nodes as the one in (13). The same is true for the other examples. But (14) and (15) are standard representations of movement/chain formation, where in (14) \( A \) and in (15) \( B \) has moved. In other words, since (13), (14) and (15) are exactly the same tree, Move or Chain is undefinable in the syntactic system that dispenses with the locality condition of the ID relation. (16) and (17) are corresponding 'remerge' structures, that indicate that the conclusion carries over to these. See Brody 2005a for a discussion of this weaker definition of syntactic trees.

8. Summary

I argued for a symmetric syntax that uses unordered (labelless) sets, where each node is understood as the set of all nodes (non-reflexively) dominated by it. Chain construction is interpretive and free ("link alpha"), subject to constraints like supersalience and checking. (Bare) checking (doing much of the work of the movement rules of earlier theories) is also symmetric. There is a single kind of linear ordering provided by the domination relation that captures top to bottom antisymmetry of the tree and which is exploited by interpretive labelling. This is the common basis of both (morpho-) phonological linearization and (quasi-) semantic c-command phenomena.

References

http://ling.auf.net/lingBuzz/000195
Pesetsky, David and Esther Torrego (2004). The syntax of valuation and the interpretability of features. Ms. MIT and Umass/Boston