Against successive cyclicity: A proof-theoretic account of extraction pathway marking

Yusuke Kubota[1] *and Robert Levine*[2]
[1] NINJAL
[2] Ohio State University

ABSTRACT

This paper proposes a novel analysis of extraction pathway marking in Type-Logical Grammar, taking advantage of proof-theoretic properties of logical proofs whose empirical application has so far been underexplored. The key idea is to allow certain linguistic expressions to be sensitive to the intermediate status of a syntactic proof. The relevant conditions can be stated concisely as constraints at the level of the proof term language, formally a special type of λ-calculus. The proposed analysis does not have any direct analog to either of the two familiar techniques for analyzing extraction pathway marking, namely, successive cyclic movement in derivational syntax and the SLASH feature percolation in HPSG.

Moreover, the 'meaning-centered' perspective that naturally emerges from this new analysis is conceptually revealing: on this approach, extraction pathway marking essentially boils down to a strategy that certain languages employ to overtly flag the existence of a semantic variable inside a partially derived linguistic expression whose interpretation is dependent on a higher-order operator that is located in a larger structure.

Keywords: long-distance dependencies, extraction pathway marking, successive cyclicity, phase, locality, Type-Logical Grammar, proof theory

1 INTRODUCTION

A widely entertained assumption in generative syntax holds that the long-distance movement operation is 'successively cyclic' (Chomsky 1973, 1977). This assumption is a fundamental part of the theory in virtually all avatars of derivational syntax since the 1970s, and is standardly taken to constitute an explanation for why movement operations in natural language are constrained in the way they appear to be, reflected in phenomena such as island constraints (see Section 2 for more on this).[1] The status of islands has been questioned much in the recent literature, but successive cyclicity is taken to receive more direct empirical evidence from typologically diverse languages in the so-called *extraction pathway marking* (EPM) phenomena (Kayne and Pollock 1978; McCloskey 1979; Chung 1982; Zaenen 1983; Borsley 2010; van Urk and Richards 2015, among others). In EPM, a syntactically displaced expression (such as the fronted *wh*-phrase in *wh*-questions)[2] induces overtly visible effects at the intermediate landing sites of a chain of movement linking the filler and the gap.

This can be illustrated most clearly by the choice of complementizer in Irish reported in McCloskey 1979. For expository convenience, we illustrate the pattern by a pseudo-language called Iringlish, which is like Irish in having the relevant distinction of two complementizers but is identical to English in all other respects.[3] As shown in (1), Iringlish (or Irish) has two complementizers *aL* and *goN* that are in complementary distribution: *aL* is used when the complementizer position

[1] See Pullum (1992) for an insightful and critical survey of the theoretical status of 'transformational cycle' in the history of generative grammar.

[2] Note that the 'filler' is not always overt, as in the case of zero relatives in English *the book I thought John read* __.

[3] The complementary distribution of the *a*- and *g*- series of Irish complementizers has been extensively discussed in James McCloskey's work (see, e.g., McCloskey 1979, 1990, 2002); for an alternative view of the morphosyntactic status of these markers, see Sells (1984). We follow McCloskey's notation in his use of upper-case letters to identify the lenition- and nasalization-triggering effects of these markers as part of the Irish Gaelic mutation system; for a recent overview of this pattern across the Celtic languages, see Iosad 2023.

is crossed by *wh*-movement (here, the covert movement of the relativization operator *Op*); *goN* appears elsewhere.

(1) a. the man *Op* **aL** [I said __ **aL** [I thought __ **aL** [__ would be there]]]

b. the man *Op* **aL** [he said __ **aL** [__ thought **goN** [he would be there]]]

The goal of the present paper is to propose an alternative account of extraction pathway marking in a proof-theoretic variant of categorial grammar (CG) known as Type-Logical Grammar (TLG). Detailed analyses of EPM effects are currently lacking in the CG literature.[4] The analysis we argue for is novel in that it does not recognize either successive cyclic movement or feature percolation of the sort utilized in the non-movement analyses of extraction pathway marking (Bouma *et al.* 2001). This surprising result comes from trying to analyze this phenomenon in a theory in which neither device is native to the underlying architecture.

The new analysis we advocate capitalizes on the proof-theoretic perspective inherent to TLG, but its core idea is arguably more general and has clear connections to the leading ideas behind many proposals within mainstream syntax (at least at an abstract level). The key claim of the present paper is that extraction pathway marking can be best understood as a 'strategy' that the grammar of some languages employs in making the intermediate (or 'incomplete') status of linguistic composition (formalized as proofs in TLG) visible in surface syntax. Making direct reference to the structure of proofs is a controversial move within the linguistic tradition of TLG (or categorial grammar research more generally). We argue that this is precisely what is needed to account for extraction pathway marking, and that by making this move, we gain conceptual clarity: the proof-theoretic perspective *predicts* the existence of extraction pathway marking in natural language, in the sense that the phenomenon exploits exactly what the grammar offers as available resource, in a conceptually simple way.

[4] The only exception we are aware of is Kubota and Levine (2020), which – as the authors themselves admit – is essentially a clumsy rendering of the HPSG-style feature percolation analysis by Bouma *et al.* (2001) within TLG.

We believe that this somewhat contentious claim would be of interest to many syntacticians and semanticists, in both 'mainstream' and 'non-mainstream' approaches. To cater to different types of audience with different backgrounds, the presentation of the material in what follows is somewhat nonstandard: after reviewing the history of the notion of cyclicity in mainstream syntax in Section 2, we present the key component of the analysis in informal terms in Section 3. This is followed by a self-contained quick review of TLG in Section 4. Section 5 then presents the analysis in full detail (Section 5.1), and puts it into perspective in relation to three larger issues: Section 5.2 examines a wider range of languages and addresses a recent claim by van Urk and Richards (2015) and van Urk (2020), according to which *both* the movement-type mechanism and the feature percolation-type mechanism are needed for a proper analysis of EPM; Section 5.3 briefly discusses implications for other phenomena pertaining to cyclicity such as reconstruction effects; Section 5.4 offers a brief comparison with a feature percolation analysis in HPSG. Section 6 concludes the paper.

2 THE STATUS OF THE NOTION OF CYCLICITY IN DERIVATIONAL SYNTAX

In this section, we review the theoretical background on the notion of cyclicity (Section 2.1) and the empirical literature on extraction pathway marking (Section 2.2). The empirical and theoretical literature is entangled in a quite complex manner, as this topic directly pertains to one of the core issues in modern syntax: the proper characterization of long-distance dependencies in natural language. The main points we aim to establish in this section are the following:

(i) The notion of cyclicity is standardly taken to constitute a fundamental principle from which various 'locality' conditions (such as island sensitivity) are supposed to follow, but this syntax-oriented perspective has come under increasing scrutiny over the years.

(ii) Many of the reported cases of alleged 'evidence' for EPM/cyclicity are also controversial since they are often based on incorrect empirical generalizations or lack proper comparison with alternative analyses that don't rely on cyclicity.

It should be noted at the outset that by making these critical remarks on the previous syntactic literature, we do not mean to claim that there is nothing that needs to be encoded in syntax to account for the EPM patterns. Rather, our point is merely that the notion of cyclicity merits reconceptualization, and that empirical evidence for it should be scrutinized at the same time in such critical rethinking. We argue that the semantically-oriented reconceptualization we propose in Section 3 (and demonstrate further in Section 5) does offer a new perspective on the relevant empirical facts themselves, by identifying this phenomenon as an overt manifestation of the intermediate status of linguistic composition of 'variable-containing' expressions.

A brief history 2.1

The notion of cyclicity as the basis for long-distance dependencies has its origins in Chomsky's (1973) proposal to derive Ross's (1967) Complex Noun Phrase Constraint (CNPC) from more general principles. Chomsky specified certain syntactic positions, specifically S and NP, as *bounding nodes* and stipulated that no more than one of such bounding nodes could be crossed at a time. Further extensions of this perspective in Chomsky (1981, 1986) led to the so-called 'Barriers' model, in which the configurational restrictions on movement were made to follow from the distinction between constituents which are 'lexically selected' and those which are not. But irrespective of precisely how the configurational restrictions on extraction were defined, the fundamental basis for such restrictions has always been entangled with the key premise in Chomsky (1973) that long-distance dependencies are an epiphenomenon of local movements chained together through unbounded iterations, and that restrictions on such dependencies are due to syntactic conditions which break such cyclically created chains.

From the early days on, it has been recognized that the mere compatibility of the distribution of islands with one or another set of syntactic configurations does not on its own amount to positive evidence for some particular set of principles of the sort Chomsky proposed. For this reason, the discovery of morphosyntactic or phonological effects that mirror the pattern of cyclic movement via bounding nodes

was important. Such 'syntactic reflexes' of cyclicity have been called extraction pathway marking (EPM) effects. See Clements *et al.* 1983 and Zaenen 1983 for earliest theoretical discussions. Reported cases of EPM in the early literature include complementizer choice in Irish (McCloskey 1979), subject-auxiliary inversion in French (Kayne and Pollock 1978) and verb agreement in Chamorro (Chung 1982).

Although the underlying architecture of the derivational theory has changed significantly over the years, especially after the advent of the Minimalist Program (MP), the idea behind cyclic movement has essentially survived to date. In the MP formulation, the notion of 'phase' – a syntactic domain where the complement of the functional head is transferred to PF at certain points in the derivation – has technically replaced the older variants of the idea of cyclic movement through certain syntactically designed positions.

Just as the main motivation of Chomsky's (1973) original proposal was to reduce some of the island effects to more general notions, the main theoretical import of the notion of phase is understood to lie in the fact that it serves as the underlying principle from which superficially observable phenomena such as island effects are to be derived. And just as in the Transformational era, the EPM effects continue to be regarded as major empirical evidence. But the status of the notion of cyclicity has constantly been controversial. Importantly, this controversy includes explicitly skeptical views within the Minimalist literature itself on attempts to derive islandhood from phasehood. First we briefly review two such remarks below. This is followed by a critical review of some of the alleged major evidence for EPM.

In a series of papers culminating in his short monograph (Boeckx 2012), Cedric Boeckx argues – building on unpublished work by Markéta Ceplová – that essentially no version of phases will actually wind up defining islands. As an example, Boeckx (2012) considers the attempt by Müller (2010) to derive Huang's (1982) Condition on Extraction Domains from the Phase Impenetrability Condition. The main conclusion of Boeckx is that Müller's attempt fails: a certain set of assumptions about constraints on feature checking and Merge that make crucial reference to the lexical valence list of heads have the unintended consequence that a *wh*-word can escape the boundary created by phase and move to a higher position (see Boeckx 2012, 63–71, and Kubota and Levine 2020, 284–289 for more details).

Den Dikken (2018) arrives at a similar conclusion, from a somewhat different angle. Following the treatment of valuation in Epstein and Seely 2002, den Dikken points out that on that analysis, information about material that is supposedly buried deeply within successive layers of phases must still be retained (i.e., made visible) to the end of the derivation. This leads him to conclude that matrix C should have access to that information, 'which should enable it to attract [a *wh*-word] straight to its specifier, without any intermediate stopovers being necessary along the way' (den Dikken 2018, 65–66). The point here is that the Epstein/Seely formulation embodies an inherent dilemma: the non-local access of information allowed for matrix C would effectively nullify the locality constraint that the very notion of phase/cyclicity is supposed to capture.

If these authors are right, we cannot automatically assume the long-held idea that the notion of phasehood is partly motivated by the explanatory role that it plays in deriving islandhood.[5] This then means that the role that the empirical phenomenon of EPM plays in motivating the theoretical notion of phase and cyclicity is now even bigger than before. This in turn motivates the central goal of the present paper, namely, looking at this notion from a different theoretical angle, one that has logical inference for meaning composition at its core. But before getting to that point, we need to critically review the alleged empirical evidence for cyclicity/EPM, since this empirical literature itself also merits careful scrutiny.

2.2 *Empirical issues*

There is now a vast literature on reported cases of empirical evidence for EPM. See, for example, van Urk's (2020) recent survey. However,

[5] Also relevant here is the fact that there is now a growing body of literature providing alternative, pragmatic or processing-oriented accounts of many of the classical island constraints. Some important work in this strand of research includes Deane (1992), Kluender (1992, 1998), Hofmeister and Sag (2010) and Chaves and Putnam (2020). See Newmeyer 2016 and Kubota and Levine 2020, Chapter 10 for recent overviews. Even within the Minimalist literature, some authors, such as Dennis Ott, go so far as to claim that islandhood is an 'open wound of syntactic theorizing' (Ott 2014, 290).

upon closer scrutiny, it turns out that there is far more room for dispute than is generally acknowledged. In another recent survey, den Dikken (2018, 69) even goes on to note that '[t]he vast majority of the arguments for successive-cyclic movement available in the literature are based on facts that are at best merely compatible with the hypothesis, not evidence for it'. In this subsection, we review some important counterarguments (some of which seem to have been underestimated) to some of the well-known cases of EPM effects.

2.2.1 EPM effects in French and Chamorro

Among the original group of languages singled out as reflecting EPM effects, French and Chamorro have come in for significant challenge. In the case of French, the acceptability of some of the key examples from Kayne and Pollock (1978) that supposedly demonstrate subject inversion in structurally higher clauses by extraction from a finite embedded clause has been called into question by Bonami *et al.* (1999); according to the latter authors, in such cases only the subject of an embedded clause projected from a head hosting the gap site can undergo this kind of inversion. On the basis of this observation and a wider range of data, Bonami *et al.* argue for an alternative analysis in which the inversion of the subject reflects generalizations about word order rather than sensitivity of an extraction pathway.

In the case of Chamorro, in Chung's (1982) original account, verbs register an agreement pattern with an argument that contains a gap, no matter how deeply embedded. However, even setting aside the theoretical problems (see den Dikken 2017), this account has an empirical flaw: the characterization of the phenomenon by Chung has been argued by Donohue and Maclachlan (1999) to be compatible with an alternative analysis that doesn't rely on the notion of cyclic movement. On the latter authors' view, in what they label 'Philippine-type languages', erosion of a typologically general pattern of voice marking has created the illusion of an exclusive agreement relationship between arguments containing gap sites and the selecting verb.

2.2.2 'Remnant movement' in Afrikaans

The earliest argument for EPM based on partial *wh*-movement, which is essentially a special case of remnant movement, comes from du Plessis (1977), with the paradigm given in (2).

(2) a. Waarvoor werk ons nou eintlik __?
wherefore work we now actually
'For what do we actually work?'

b. Waar werk ons nou eintlik __ **voor**?

c. Waarvoor [dink julle __ [werk ons __]]?
'What do you think we work for?'

d. Waar/wat dink julle [**voor** __ [werk ons __]] ?

(2a) exhibits the more or less default extraction pattern: *waarvoor* appears in Spec,CP with a gap in its presumed argument position. In (2b), however, *waar* has moved, but has left behind the bound form of the preposition with which it is compounded in (2a). (2c) is a long-distance pattern of full *waarvoor* extraction, and (2d) is the crucial case in which *voor* is stranded at an intermediate Spec,CP.

However, as discussed in den Besten (2010), the interpretation of the facts just given appears to be simply mistaken, or at least equivocal (see also den Dikken (2009), who refutes similar arguments for cyclicity in Dutch based on similar sorts of considerations).[6] In particular, den Besten notes that in (2d), the application of the matrix V2 rule in Afrikaans (moving the verb *dink* from the clause-final underlying position immediately before the complement clause to the surface position) makes it difficult to tell whether *voor* actually occupies the embedded Spec,CP position or is an element of the matrix clause syntactically. Since Afrikaans V2 is a root clause constraint, one can observe *voor*'s actual underlying location more accurately using an embedded *wh*-interrogative example:

[6] There is another problem with this remnant movement analysis. As noted by du Plessis himself, the alleged stranded preposition in (2b) and (2d) has to be *voor*, instead of the standard free form preposition *vir* (as in *vir wat* 'for what'). *Voor* is identical in form with the part of the compound *wh*-PP *waarvoor* in (2a), and this form identity is supposedly what motivates du Plessis' analysis via remnant movement. However, outside of this remnant movement literature, there is no known case in which a syntactic operation pries apart a lexical item in the way it does in (2b) and (2d) (on du Plessis' analysis). It is unclear how this violation of the lexical integrity principle (see, e.g., Bresnan and Mchombo 1995; Manning *et al.* 1999) can be accounted for in a cyclic movement-based analysis.

(3) a. Ek sou graag wou weet [$_{CP}$ waar [julle **voor** dink [$_{CP}$ dat [ons werk]]]]
'I would like to know what you think we work for.'

b. *Ek sou graag wou weet [$_{CP}$ waar [julle dink [$_{CP}$ **voor** [dat ons werk]]]]

The contrast in (3) shows that *voor* can end up stranded as a matrix clause element (presumably via clause-internal fronting of the *wh*-element) but cannot occupy an embedded Spec,CP. According to den Besten (2010), du Plessis's (1977) crucial example (2d) should thus be analyzed on a par with (3a) (modulo the V2 word order) rather than the ungrammatical (3b), and hence cannot be taken to involve an intermediate Spec,CP remnant.

2.2.3 *Wh*-copying

The *wh*-copying construction in German (and some other languages) has often been invoked in the literature as evidence for successive cyclicity. This phenomenon is illustrated in (4), where a copy of the *wh*-word appears in overt syntax at an intermediate Spec,CP position:

(4) **Wen** meint Karl [$_{CP}$ **wen** wir __ gewählt haben]?
who thinks Karl who we voted.for have
'Who does Karl think we voted for?'

Den Dikken (2017) notes several issues with an analysis of *wh*-copying in terms of successive cyclic movement. First, as den Dikken notes, prospects for a cyclic analysis start looking murky as soon as we turn our attention to cases involving complex *wh*-phrases.

(5) a. ***Wessen Studenten** denkst du **wessen Studenten** man
whose students think you whose students one
einladen sollte?
invite should
intended: 'Whose students do you think should be invited?'

b. **Wen** denkst du [**wen von den Studenten**] man
who think you who of the students one
einladen sollte?
invite should
'Which of the students do you think should be invited?'

(5a) shows that pronouncing a literal copy of a complex *wh*-phrase at each landing site is ungrammatical. The example improves by replacing one of the two complex *wh*-phrases by a simpler form as in (5b). This is exactly the opposite of what one would expect on the simplest version of 'form-identical multiple copy'-type analysis.

The above paradigm seems already quite troublesome, but den Dikken notes further difficulties for a cyclic movement analysis. Specifically, with respect to scope interpretation, the *wh*-copying phenomenon does not behave like standard overt long-distance movement, but is more similar to the *wh*-scope marking construction (e.g., *Was meint Karl wen wir __ gewählt haben?*, where instead of the *wh*-pronoun *wen*, the *wh*-word at the matrix level is the fixed form *was* 'what'). This and the problem with complex *wh*-phrases leads den Dikken to conclude that the *wh*-copying construction had better be analyzed as a special type of *wh*-scope marking and should not be viewed as a case of long-distance movement with copies in a single derivational chain pronounced at intermediate and final landing sites.

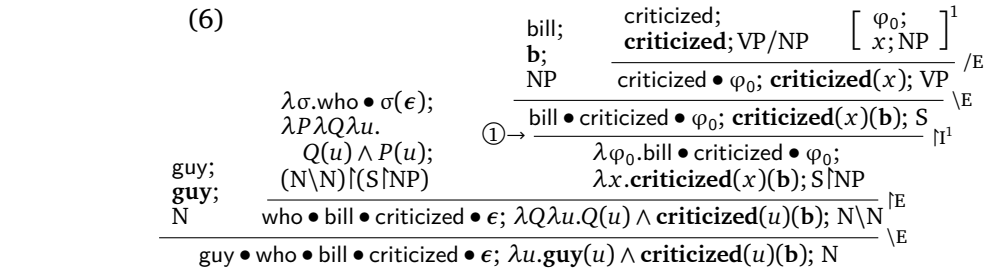
SKETCH OF A NEW ANALYSIS 3

A characteristic that distinguishes our approach from all known formulations of cyclicity in the literature is that it takes the cyclicity effect to be a reflex of the way in which meaning composition interacts with syntax. This is technically implemented via constraints on the forms of logical proofs corresponding to linguistic derivations. The full formal analysis (presented in Section 5) is formulated in a version of Type-Logical Grammar (TLG), whose formal details may feel dauntingly technical to some. However, as explained below, it can essentially be seen as a formalization of the LF-based theory in mainstream syntax. To make the exposition easier to follow, we present the analysis in two steps. This section presents the gist of the analysis in informal terms. This is followed by a compact introduction to TLG in Section 4 and the full formal analysis of EPM in Section 5.

3.1 *Derivations as proofs*

In TLG, linguistic derivations are formally logical proofs. Roughly speaking, Merge (in minimalist terms) corresponds to modus ponens ($P \to Q, P \models Q$) and Move to hypothetical proof (assuming P, deriving some conclusion Q, and then, drawing the real conclusion $P \to Q$ by withdrawing the hypothesis P). The following derivation for the relative clause *who Bill criticized* __ illustrates the relevant point:

(6)



Here, linguistic signs are written as triples of prosodic form, semantics and syntactic category (or 'syntactic type'). The key steps in the derivation in (6) can be paraphrased in prose as follows.

- The NP with prosody φ_0 is a hypothetically assumed NP (the square brackets around it indicate its status as such). With this hypothesis, we derive a complete S corresponding to the body of the relative clause *Bill criticized* __ (immediately above ①).
- The crucial step is the next one (①). At this point, the hypothesis is *withdrawn*, yielding an expression of category S↑NP, a sentence containing an NP-type gap.
- The relative pronoun then takes this gapped sentence as its first argument and returns a backward nominal modifier of type N\N.

The exact way in which prosodic lambda binding in (6) ensures the effect of 'overt movement' of the relative pronoun will be discussed in Section 4, so, we omit the details here. The important point here, which will be crucial in the implementation of EPM, is that hypothetical reasoning (deriving a gapped S↑NP from a *hypothetical* proof of S on the assumption of NP) is the underlying principle that derives the effect of 'movement' (in the standard parlance) and that syntactic/prosodic form and semantics are derived in tandem at each step.

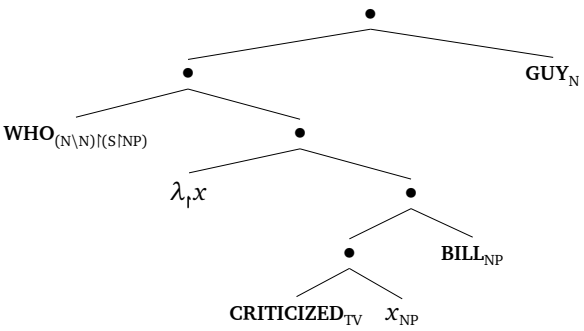
To facilitate the ensuing discussion, we notate the proof trees of the sort in (6) in an alternative, simpler format. Again, we gloss over details radically in this section. All one needs to know at this point is that this alternative notation has solid theoretical underpinnings (explained in detail in Section 4) and that it looks very similar to LF trees of the sort familiar from, e.g., Heim and Kratzer 1998.

We first posit the following constants (written in small capitals) for each of the lexical items used in the derivation in (6) (in what follows, TV is an abbreviation for (NP\S)/NP):

(7) $\textbf{CRITICIZED}_{\text{TV}}$ = criticized; **criticized**; TV
$\textbf{WHO}_{(\text{N}\backslash\text{N})\uparrow(\text{S}\uparrow\text{NP})}$ =
$\lambda\sigma.\text{who} \bullet \sigma(\epsilon)$; $\lambda P\lambda Q\lambda u.Q(u)\wedge P(u)$; $(\text{N}\backslash\text{N})\uparrow(\text{S}\uparrow\text{NP})$
$\textbf{BILL}_{\text{NP}}$ = bill; **b**; NP
$\textbf{GUY}_{\text{N}}$ = guy; **guy**; N

Then, the proof tree in (6) can be rewritten as in (8):

(8)



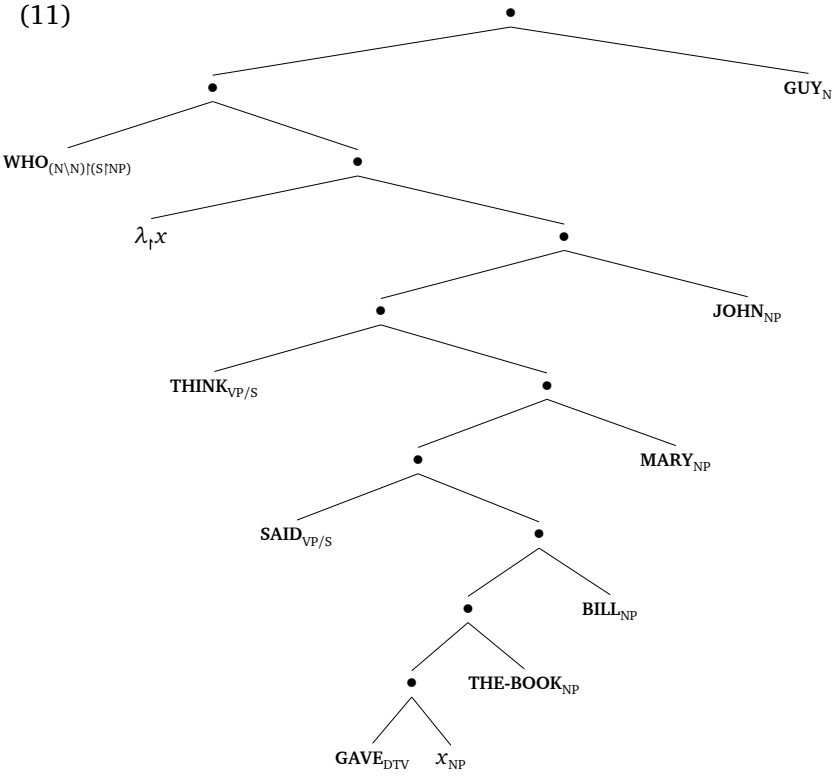
All we have done here is replace the tripartite signs at the leaves by the abbreviations in (7) and write the tree upside down. Thus, from (8) and (7), the original proof in (6) (with more information explicitly written at each node) is fully recoverable.

Note that this way of establishing the relationship between the *wh*-operator and the corresponding hypothesis can cross multiple levels of embedding, since all that's involved is the general mechanism for deducing expressions of type S⇑NP, a sentence missing an NP in some arbitrary position inside. Thus, a long-distance relativization example (9) can be analyzed by exactly the same mechanism as in (10)/(11).

(9) the guy who John thinks Mary said Bill gave __ the book

(10) $\mathbf{WHO}_{(N\backslash N)\uparrow(S\uparrow NP)}(\lambda_{\uparrow}x.\mathbf{THINK}_{VP/S}(\mathbf{SAID}_{VP/S}(\mathbf{GAVE}_{DTV}(x_{NP})$
$(\mathbf{THE\text{-}BOOK}_{NP})(\mathbf{BILL}_{NP}))(\mathbf{MARY}_{NP}))(\mathbf{JOHN}_{NP}))(\mathbf{GUY}_{N})$

(11)



3.2 *Irish complementizer marking*

We illustrate the analysis with the Irish complementizer choice reported in McCloskey 1979.[7] In this subsection, we review the key data, using our pseudo-language Iringlish from Section 1 for expository convenience. We start with clausal embedding without any extraction. In

[7] As noted by Chaves and Putnam (2020), McCloskey's original proposal in terms of cyclic movement does not seem to be entirely unproblematic in view of the Minimalist theory of movement. In the latter, movement is driven by the need to check uninterpretable features, and in McCloskey 2002, McCloskey himself is essentially forced to posit a number of uninterpretable features which themselves lack independent empirical support.

this case, as shown in (12)–(13), the complementizers (the counterpart of *that* in English) are all realized as *goN*.

(12) I thought **goN** [he would be there].

(13) I said **goN** [I thought **goN** [he would be there]].

As explained in Section 1, when the complementizer position is on an extraction pathway, the alternative form *aL* is used. Thus, for example, in the following (14), the lower clause is marked by *goN*, but the higher clause is marked by *aL*:

(14) the man **aL** [__ thought **goN** [he would be there]]

The examples in (15)–(16), with a multiple chain of *aL* complementation, show that the linkage between the filler and the gap is registered over an arbitrary number of structural levels.

(15) the man **aL** [I thought **aL** [__ would be there]]

(16) the man **aL** [I said **aL** [I thought **aL** [__ would be there]]]

Regardless of the depth of the extraction, as soon as the gap site is identified, all lower clauses which themselves are not associated with an extraction will be marked by *goN*, a point illustrated in (14) and at still greater structural depth in (17).

(17) the man **aL** [he said **aL** [__ thought **goN** [he would be there]]]

Accounting for extraction pathway marking 3.3

The pattern displayed by Iringlish is simple: the form of the complementizer is sensitive to the existence of an unbound gap in the complement clause. But how can we encode this restriction? The apparent dilemma here is that neither cyclic movement nor feature percolation is native to the architecture of TLG. In the analysis of extraction sketched above in Section 3.1, the filler/gap identification is mediated via a single instance of hypothetical reasoning. So, nothing 'moves' literally (let alone in a successive cyclic way), nor is there any structure-manipulation operation or feature percolation of any sort.

The answer comes from seeing proofs as structured objects that linguistic signs can (at least partly) make reference to. Mainstream

syntacticians will probably consider this idea more or less unobjectionable (since LF trees are representational objects anyway), but advocates of (traditional) categorial grammar may find it alarming. This is because we need to part with one influential assumption that has dominated CG research over the past several decades. What we need to give up is the idea that the grammar cannot access the internal structures of syntactic proofs.[8] For ardent advocates of direct compositionality, this may appear to be a high price to pay. For such readers, we note that the challenge here is to come up with an explicit analysis of EPM facts in a theory that abides by direct compositionality – a task which, so far as we can tell, is far from trivial.

The proof term notation of derivations introduced above enables a concise formulation of the EPM patterns exhibited by the Iringlish (or Irish) data above. We illustrate this point with a fragment of Iringlish with the lexicon in (18).[9]

(18) a. $\mathbf{WBT}_{\mathrm{NP}\backslash\mathrm{S}}$ = would • be • there; $\lambda x.\mathbf{located}(x)(\mathbf{there})$; $\mathrm{NP}\backslash\mathrm{S}$

b. $\mathbf{MAN}_{\mathrm{N}}$ = man; **man**; N

c. $\mathbf{THOUGHT}_{(\mathrm{NP}\backslash\mathrm{S})/\mathrm{S}'}$ = thought; **thought**; $(\mathrm{NP}\backslash\mathrm{S})/\mathrm{S}'$

[8] While the origin of this idea is unclear, it likely stems from the view in classical Montague Grammar that the translation language is an intermediate step that is in principle eliminable (see, e.g., Dowty *et al.* 1981 and Cooper 1983). It is worth noting in this connection that Dowty (2007), in his later work, has emphasized that compositionality is a *methodological* principle rather than a fixed or fundamental assumption.

[9] For expository convenience, the fragment presented in the main text involves an empty relativizer **REL**. Proponents of lexicalist theories of syntax might find this treatment objectionable. For such readers, we'd like to point out that the effect of **REL** can be lexicalized easily with an alternative, relativized version of *aL* shown below, which can be thought of as a lexicalization of function composition of complementizer *aL* and relativizer **REL** in the underlying calculus, as in (ia):

(i) a. **AL-REL**

$= \lambda_{\restriction} f.\mathbf{REL}(\lambda_{\restriction} x.\mathbf{AL}(f_{\mathrm{S}\uparrow\mathrm{NP}_{+wh}}(x_{\mathrm{NP}_{+wh}})))$

$= \lambda\sigma.\mathrm{aL} \bullet \sigma(\epsilon);\ \lambda P\lambda Q\lambda y.Q(y) \wedge P(y);\ (\mathrm{N}\backslash\mathrm{N})\uparrow(\mathrm{S}\uparrow\mathrm{NP}_{+wh})$

b. ***GON-REL**

$= \lambda_{\restriction} f.\mathbf{REL}(\lambda_{\restriction} x.\mathbf{GON}(f_{\mathrm{S}\uparrow\mathrm{NP}_{+wh}}(x_{\mathrm{NP}_{+wh}})))$

Note that such a lexicalized variant is unavailable for *goN*: as in (ib), it violates the free-variable prohibition restriction imposed on *goN* in (18f).

d. $\textbf{SAID}_{(\mathrm{NP}\backslash\mathrm{S})/\mathrm{S}'}$ = said; **said**; $(\mathrm{NP}\backslash\mathrm{S})/\mathrm{S}'$

e. $\textbf{AL}_{\mathrm{S}'/\mathrm{S}}$ = aL; $\lambda p.p$; S'/S
where for any α, $\textbf{AL}(\alpha)$ is defined only if $fv_{\mathrm{X}_{+wh}}(\alpha) \neq \emptyset$

f. $\textbf{GON}_{\mathrm{S}'/\mathrm{S}}$ = goN; $\lambda p.p$; S'/S
where for any α, $\textbf{GON}(\alpha)$ is defined only if $fv_{\mathrm{X}_{+wh}}(\alpha) = \emptyset$

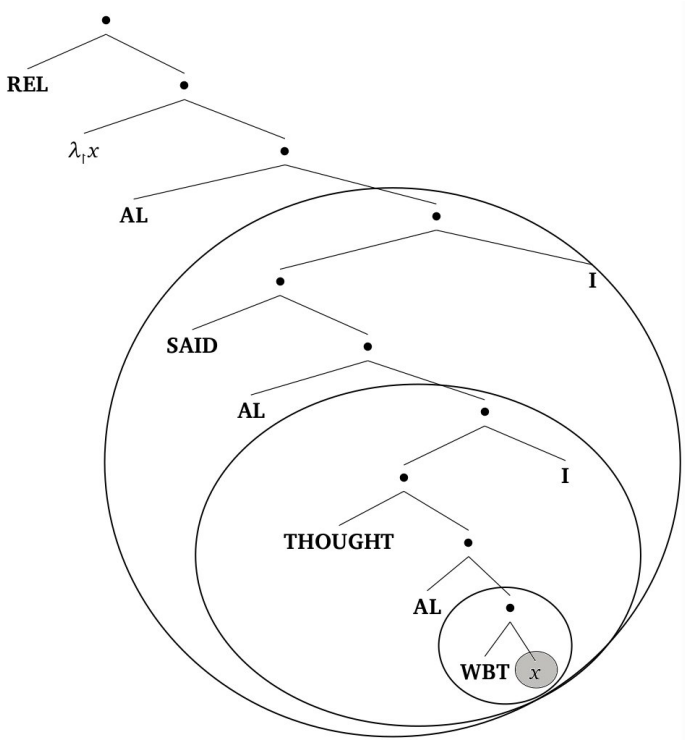
g. $\textbf{REL}_{(\mathrm{N}\backslash\mathrm{N})\restriction(\mathrm{S}'\restriction\mathrm{NP}_{+wh})}$ =
$\lambda\sigma_2.\sigma_2(\epsilon)$; $\lambda P\lambda Q\lambda y.Q(y) \wedge P(y)$; $(\mathrm{N}\backslash\mathrm{N})\restriction(\mathrm{S}'\restriction\mathrm{NP}_{+wh})$

The key components of this analysis are the restrictions imposed on *aL* and *goN* that refer to the structures of the terms given as their (first) arguments. fv_{Φ} is the standard, inductively defined function that returns all free variables contained in a term, except that it filters the output of the general purpose fv to type Φ. We illustrate with concrete examples below how these lexical constraints on complementizers properly restrict their distributions.

The topmost relative clause in (19) can now be derived as in (20).

(19) the man **aL** [I said **aL** [I thought **aL** [__ would be there]]]

(20)



Here, each token of *aL* applies to a clausal complement containing the free variable x and hence is legal.

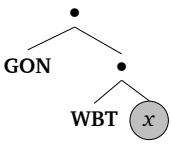
The ungrammaticality of the examples in (21) also follows immediately. In the case of (21a), *goN* is used instead of *aL* in the subproof corresponding to the innermost clause. This violates the constraint $fv_{\mathrm{X}_{+wh}}(\alpha) = \emptyset$ on the first argument of *goN*. Similarly, in (21b), *goN* replaces the first *aL* in the subproof corresponding to the outermost clause. Here again, the relevant 'no unbound $+wh$ hypothesis' constraint on *goN* is violated.

(21) a. *the man **aL** [I said **aL** [I thought **goN** [__ would be there]]]

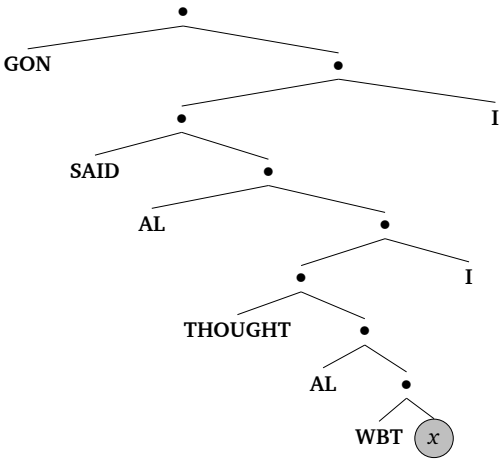
b. *the man **goN** [I said **aL** [I thought **aL** [__ would be there]]]

The offending subterms in the proofs for (21a,b) are shown in (22).

(22) a.



b.



Thus, by making the lexical entries of the complementizers sensitive to the existence of open hypotheses in subproofs, we obtain a simple and straightforward analysis of EPM. Since the existence of open hypotheses conceptually corresponds to the fact that the complementizer is licensed at a point in the derivation at which filler-gap linkage

is not yet established, we obtain the effect of 'cyclicity' without literally encoding a structure-manipulation operation of cyclic movement.

Some remarks are in order regarding the possible similarities and differences between the present analysis of EPM as 'proof structure making' and the more standard configurational approach in derivational syntax (see Citko 2014 for an overview of the latter). The similarity should be clear. In both approaches, linguistic derivations are regarded as structured objects and the grammar offers one way or another for making reference to part of the 'derivational history' that certain lexical items (or other aspects of grammar) are sensitive to.

Turning to differences, we see at least two aspects in which our proposal substantially differs from the standard view. First, by viewing EPM as a mere reflection of the 'hypothesis containing' status of a subproof, our approach predicts that 'phase boundaries' are not necessarily limited to a small set of categories (standardly, CP and vP).[10] This is perhaps the single most important difference. What constitutes the exact set of 'phase boundaries' is itself a controversial issue in Minimalist syntax (see Legate 1998 and especially Matushansky 2005 for some discussion on this thorny issue), and we are not prepared to get into an in-depth discussion on this topic, but one point is worth noting: in Minimalist formulations, there has to be some conceptual basis for restricting the set of 'phase boundaries', and it has sometimes been suggested that this may come from semantic considerations, with CP corresponding to a proposition-denoting unit (cf., e.g., Chomsky 2000; Hinzen 2012). If such a semantic characterization of 'phase boundaries' is tenable, that would be entirely compatible with our account, since in TLG, there is a tight correspondence between syntactic types and semantic types, and at each step of derivation, the full denotation of the linguistic expression being derived is available.

This then relates to the second major difference. In the standard phase-based approach, the correspondence between syntactic computation and compositional semantics is somewhat unclear. It is only via the explicit structural operation of movement (or external merge and

[10] In Section 6, we offer brief speculations on how one might go about making sense of what seems like a skewed syntactic distribution of EPM items cross-linguistically under the meaning-centered approach that our proof-theoretic perspective embodies.

the specific way in which two copies of the same lexical item get interpreted at the CI component) that we get the effect of variable binding. Our approach captures the connection between 'movement' and 'variable binding' more straightforwardly, since 'movement' is by definition nothing other than variable binding (or hypothetical reasoning) in the underlying logic governing the correspondence between surface form and the compositional meaning. The analysis of EPM crucially exploits this property of the TLG architecture (and the formal tools available in it for formulating meta-statements pertaining to the statuses of subproofs), a point we get back to at the end of the paper.

Our approach essentially embodies a *meaning-centered* perspective on EPM. We believe that this represents at least an interesting enough alternative to the standard structure-driven approach. It may appear to have some glaring loose ends, but we believe that the conceptual simplicity is attractive enough to compensate for this possible shortcoming (which after all relates to a still open and controversial issue).

4 LONG-DISTANCE DEPENDENCIES IN HYBRID TLG

This section is meant to serve two purposes: to introduce Hybrid TLG as a syntactic framework and to illustrate its workings with an analysis of pied-piping in relative clauses. The choice of the empirical phenomenon is motivated by the fact that pied-piping exhibits properties of both 'overt' and 'covert' movement in derivational syntax. A recasting of the movement-based analysis of pied-piping from mainstream generative syntax in Hybrid TLG – building on an earlier analysis by Morrill (1994) – illustrates clearly the way in which TLG handles complex mapping between form and meaning. There is already substantial literature on linguistic applications of TLG (see, e.g., Morrill 1994; Carpenter 1997; Kubota and Levine 2020), and readers are encouraged to refer to these sources for more information about TLG as a syntactic framework. Handbook articles such as Moortgat 2011, 2014 and Kubota 2021 are also useful sources of reference.

The full system of Hybrid TLG comprises three logical connectives $/$, $\backslash$ and $\uparrow$, and has Elimination and Introduction rules for all these.

However, since the linguistic phenomena we deal with in this paper do not involve hypothetical reasoning with the directional slashes / and \, our presentation below focuses on the way in which the directional slashes / and \ are used for licensing local function-argument structures and on the use of the ↑ connective for modeling 'movement' operations (this corresponds to the system introduced in Section 2.3 of Kubota and Levine 2020). The more complex Introduction rules for / and \ are discussed only briefly in Section 4.4.

AB grammar 4.1

We start with a simple fragment called the *AB grammar* (Ajdukiewicz 1935; Bar-Hillel 1953), consisting of just the two syntactic rules in (23):

(23) a. Forward Slash Elimination

$$\frac{b;\ B \quad a;\ A/B}{a \bullet b;\ A}\ /\mathrm{E}$$

b. Backward Slash Elimination

$$\frac{b;\ B \quad a;\ B\backslash A}{b \bullet a;\ A}\ \backslash\mathrm{E}$$

With the somewhat minimal lexicon in (24), we can license a simple transitive verb sentence (25) as in (26). The two slashes / and \ are used to form complex syntactic categories, or syntactic types, indicating valence information: The transitive verb *loves* is assigned the syntactic type (NP\S)/NP since it first combines with an NP to its right (i.e. the direct object) and then another NP to its left (i.e. the subject).

(24) a. john; NP
b. mary; NP
c. ran; NP\S
d. loves; (NP\S)/NP

(25) John loves Mary.

(26)

$$\dfrac{\mathsf{john};\ \mathrm{NP} \qquad \dfrac{\mathsf{mary};\ \mathrm{NP} \quad \mathsf{loves};\ (\mathrm{NP}\backslash\mathrm{S})/\mathrm{NP}}{\mathsf{loves} \bullet \mathsf{mary};\ \mathrm{NP}\backslash\mathrm{S}}\ /\mathrm{E}}{\mathsf{john} \bullet \mathsf{loves} \bullet \mathsf{mary};\ \mathrm{S}}\ \backslash\mathrm{E}$$

There is one thing to keep in mind about proof notation. In the presentation of proofs and rules adopted in (23) and (26), the word order is reflected solely in the prosodic annotations at each node of the tree, and the left and right order of the premises in a subtree does not have anything to do with the surface word order of English sentences (in

the rest of the paper, we generally align the order of premises with the actual word order, but this is only for expository ease).

Syntactic types are defined recursively. For the AB grammar, this can be concisely written using the so-called 'BNF notation' as follows (the exact choice of the set of basic types is an empirical question):

(27) $\mathscr{A} := \{ \text{S, NP, N, PP, ...} \}$ (atomic type)
$\mathscr{T} := \mathscr{A} \mid \mathscr{T}\backslash\mathscr{T} \mid \mathscr{T}/\mathscr{T}$ (type)

In words, anything that is an atomic type is a type, and any complex expression of form $A\backslash B$ or A/B where A and B are both types is a type.

As should already be clear in the above illustration, categorial grammar lexicalizes the valence (or subcategorization) properties of linguistic expressions, and this is transparently represented in the syntactic types of functional expressions (such as verb lexical entries). Here are some more sample lexical entries:

(28) a. ran; NP\S
b. read; (NP\S)/NP
c. introduces; (NP\S)/PP/NP

4.2 *Syntax-semantics interface*

Assuming the standard recursive definition of semantic types as in (29) (with basic types e (individuals) and t (truth values) for an extensional fragment), we can define the function Sem that returns, for each syntactic type given as input, its semantic type, as in (30)–(31).

(29) a. $\mathscr{A}_\sigma := \{ e, t \}$ (atomic semantic type)
b. $\mathscr{T}_\sigma := \mathscr{A}_\sigma \mid \mathscr{T}_\sigma \rightarrow \mathscr{T}_\sigma$ (semantic type)

(30) (Base Case)
a. $\mathsf{Sem}(\text{NP}) = \mathsf{Sem}(\text{PP}) = e$
b. $\mathsf{Sem}(\text{N}) = e \rightarrow t$
c. $\mathsf{Sem}(\text{S}) = t$

(31) (Recursive Clause)
For any complex syntactic type of the form A/B (or $B\backslash A$),
$\mathsf{Sem}(A/B)$ $(= \mathsf{Sem}(B\backslash A)) = \mathsf{Sem}(B) \rightarrow \mathsf{Sem}(A)$

For example, assuming that VP adverbs such as *quickly* are of type (NP\S)\(NP\S), we can determine their semantic type based on the syntactic type by following the definitions in (29)–(31):

(32) $\mathsf{Sem}((\mathrm{NP}\backslash\mathrm{S})\backslash(\mathrm{NP}\backslash\mathrm{S}))$
$= \mathsf{Sem}(\mathrm{NP}\backslash\mathrm{S}) \rightarrow \mathsf{Sem}(\mathrm{NP}\backslash\mathrm{S})$
$= (\mathsf{Sem}(\mathrm{NP}) \rightarrow \mathsf{Sem}(\mathrm{S})) \rightarrow (\mathsf{Sem}(\mathrm{NP}) \rightarrow \mathsf{Sem}(\mathrm{S}))$
$= (e \rightarrow t) \rightarrow (e \rightarrow t)$

In other words, the syntactic type (NP\S)\(NP\S) transparently represents the semantic type of a VP modifier as an $e \rightarrow t$ property modifier.

Syntactic rules with semantics can then be written as in (33) (where the semantic effect of these rules is *function application*) and a sample derivation with semantic annotation is given in (34).

(33) a. Forward Slash Elimination

$$\frac{a;\ \mathscr{F};\ A/B \quad b;\ \mathscr{G};\ B}{a \bullet b;\ \mathscr{F}(\mathscr{G});\ A}\ /\mathrm{E}$$

b. Backward Slash Elimination

$$\frac{b;\ \mathscr{G};\ B \quad a;\ \mathscr{F};\ B\backslash A}{b \bullet a;\ \mathscr{F}(\mathscr{G});\ A}\ \backslash\mathrm{E}$$

(34)

$$\dfrac{\text{john};\ \mathbf{j};\ \mathrm{NP} \qquad \dfrac{\dfrac{\text{chased};\ \mathbf{chased};\ (\mathrm{NP}\backslash\mathrm{S})/\mathrm{NP} \quad \text{mary};\ \mathbf{m};\ \mathrm{NP}}{\text{chased} \bullet \text{mary};\ \mathbf{chased}(\mathbf{m});\ \mathrm{NP}\backslash\mathrm{S}}\ /\mathrm{E} \qquad \begin{array}{l}\text{patiently};\\ \mathbf{patiently};\\ (\mathrm{NP}\backslash\mathrm{S})\backslash(\mathrm{NP}\backslash\mathrm{S})\end{array}}{\text{chased} \bullet \text{mary} \bullet \text{patiently};\ \mathbf{patiently}(\mathbf{chased}(\mathbf{m}));\ \mathrm{NP}\backslash\mathrm{S}}\ \backslash\mathrm{E}}{\text{john} \bullet \text{chased} \bullet \text{mary} \bullet \text{patiently};\ \mathbf{patiently}(\mathbf{chased}(\mathbf{m}))(\mathbf{j});\ \mathrm{S}}\ \backslash\mathrm{E}$$

4.3 *Adding the vertical slash for 'movement'*

The AB grammar introduced above deals with local licensing of arguments via the Elimination rules for / and \. This roughly corresponds to simple phrase structure grammar (or context-free grammar) without 'movement' operations. In order to model phenomena that involve both 'covert' and 'overt' movement (in the derivational terminology), we need to extend the underlying logic. In Hybrid TLG, this is done by introducing functional expressions in the prosodic representations of linguistic signs written as λ-terms (Oehrle 1994; de Groote 2001; Muskens 2003; Mihaliček and Pollard 2012). As will become clear below, λ-binding of variables in the prosodic representations makes it possible to 'reason about' linguistic expressions in which something is missing in the middle. This technique is crucially exploited in the analysis of relative clauses in (38) and (40) below.

Building on this tradition, we introduce into our system a new connective ⇂ called the *vertical slash*, for order-insensitive mode of implication (as with /, we write the argument to the right for ⇂). For this connective, we posit the following two rules:

(35) a. Vertical Slash Introduction

$$\dfrac{\begin{array}{c}\vdots\ \underline{[\varphi;\, x;\, A]^n}\ \vdots \\ \vdots\qquad\vdots\qquad\vdots \\ b;\, \mathscr{F};\, B\end{array}}{\lambda\varphi.b;\, \lambda x.\mathscr{F};\, B{\restriction}A}\ {\restriction}\mathrm{I}^n$$

b. Vertical Slash Elimination

$$\dfrac{a;\, \mathscr{F};\, A{\restriction}B \quad b;\, \mathscr{G};\, B}{a(b);\, \mathscr{F}(\mathscr{G});\, A}\ {\restriction}\mathrm{E}$$

Of these two rules, Vertical Slash Elimination (35b) is simpler. It licenses a structure in which a linguistic expression that has functional prosody (reflected in the syntactic type $A{\restriction}B$) combines with its argument (of syntactic type B). The rule specifies that in such function-argument pairs (i.e., $A{\restriction}B$ and B), the two items are combined by function application in both semantics and prosody.

The workings of the Vertical Slash Introduction rule (35a) is somewhat more complex, but the underlying idea is simple. This rule licenses a type of proof in which some linguistic expression (the bracketed expression with index n) is hypothetically assumed to derive an intermediate conclusion (on the penultimate line with type B). The rule then licenses an expression of type $B{\restriction}A$ by withdrawing the hypothesis A. The corresponding effect in the semantic and prosodic components is λ-binding of the variables introduced by the hypothesis A. The semantic λ-binding should make obvious sense (given the analogy to movement). What's novel (for those unfamiliar with the subspecies of CG stemming from Oehrle 1994) is the λ-binding in the prosodic component. This will be illustrated with an example below in (38). The correspondence between a hypothesis and the ⇂I step at which it is withdrawn in the proof tree is kept track of by the index n, since there may be multiple such pairs within a single proof.

The way this extended system works can be best illustrated by concrete examples, so let us now examine a simple analysis of English relative clauses. The key idea is that the new rules just introduced enable us to 'reason about' linguistic expressions in which some material is missing. For example, in (36), the body of the relative clause *Bill criticized* __ is analyzed as S⇂NP, a sentence missing an NP.

(36) the guy who Bill criticized __

We posit the following entry for the relative pronoun *who* in which both the semantics and the prosody are higher-order functions.

(37) $\lambda\sigma.\mathsf{who} \bullet \sigma(\boldsymbol{\epsilon})$; $\lambda P\lambda Q\lambda u.Q(u) \wedge P(u)$; $(\mathrm{N}\backslash\mathrm{N})\uparrow(\mathrm{S}\uparrow\mathrm{NP})$

We can then license (38) for (36) (the dotted lines in (38) just show the β-reduction steps for the prosodic term, and are not part of the syntactic derivation; in what follows, VP is an abbreviation for NP\S).

(38)

$$\dfrac{\mathsf{guy};\ \mathbf{guy};\ \mathrm{N} \qquad \dfrac{\lambda\sigma[\mathsf{who} \bullet \sigma(\boldsymbol{\epsilon})](\lambda\varphi_0.\mathsf{bill} \bullet \mathsf{criticized} \bullet \varphi_0);\ \lambda Q\lambda u.Q(u) \wedge \mathbf{criticized}(u)(\mathbf{b});\ \mathrm{N}\backslash\mathrm{N}}{\mathsf{who} \bullet \lambda\varphi_0[\mathsf{bill} \bullet \mathsf{criticized} \bullet \varphi_0](\boldsymbol{\epsilon});\ \lambda Q\lambda u.Q(u) \wedge \mathbf{criticized}(u)(\mathbf{b});\ \mathrm{N}\backslash\mathrm{N}}\uparrow\mathrm{E}}{\mathsf{who} \bullet \mathsf{bill} \bullet \mathsf{criticized} \bullet \boldsymbol{\epsilon};\ \lambda Q\lambda u.Q(u) \wedge \mathbf{criticized}(u)(\mathbf{b});\ \mathrm{N}\backslash\mathrm{N}}\backslash\mathrm{E}$$

$$\dfrac{\lambda\sigma.\mathsf{who} \bullet \sigma(\boldsymbol{\epsilon});\ \lambda P\lambda Q\lambda u.Q(u) \wedge P(u);\ (\mathrm{N}\backslash\mathrm{N})\uparrow(\mathrm{S}\uparrow\mathrm{NP}) \qquad \textcircled{1}\rightarrow\ \dfrac{\dfrac{\mathsf{bill};\ \mathbf{b};\ \mathrm{NP} \qquad \dfrac{\mathsf{criticized};\ \mathbf{criticized};\ \mathrm{VP}/\mathrm{NP} \qquad \left[\begin{array}{l}\varphi_0;\\ x;\ \mathrm{NP}\end{array}\right]^1}{\mathsf{criticized} \bullet \varphi_0;\ \mathbf{criticized}(x);\ \mathrm{VP}}/\mathrm{E}}{\mathsf{bill} \bullet \mathsf{criticized} \bullet \varphi_0;\ \mathbf{criticized}(x)(\mathbf{b});\ \mathrm{S}}\backslash\mathrm{E}}{\lambda\varphi_0.\mathsf{bill} \bullet \mathsf{criticized} \bullet \varphi_0;\ \lambda x.\mathbf{criticized}(x)(\mathbf{b});\ \mathrm{S}\uparrow\mathrm{NP}}\uparrow\mathrm{I}^1}{\lambda\sigma[\mathsf{who} \bullet \sigma(\boldsymbol{\epsilon})](\lambda\varphi_0.\mathsf{bill} \bullet \mathsf{criticized} \bullet \varphi_0);\ \lambda Q\lambda u.Q(u) \wedge \mathbf{criticized}(u)(\mathbf{b});\ \mathrm{N}\backslash\mathrm{N}}\uparrow\mathrm{E}$$

The derivation in (38) can be paraphrased in prose as follows.

- The NP with prosody φ_0 is a hypothetically assumed NP (the square brackets around it indicate its status as such). With this hypothesis, we derive a complete S corresponding to the body of the relative clause *Bill criticized* __ (immediately above ①).
- The crucial step is the next one (①). At this point, the hypothesis is *withdrawn* with the $\uparrow$-Introduction rule. This yields an S$\uparrow$NP, a sentence containing an NP-type gap. The string position of the gap is kept track of by λ-binding the prosodic variable φ_0.
- The relative pronoun, with the lexical specification in (37), then takes this gapped sentence as its first argument and returns a backward nominal modifier of type N\N. (Semantically, the relative pronoun denotes an intersective modifier of two properties.)

The final step where the relative pronoun takes a gapped sentence as argument perhaps requires some comment. The key point here is that the prosodic specification of the relative pronoun in (37) is a higher-order function that combines strings in a particular way. Specifically, its first argument σ is the gapped sentence (itself a function of type

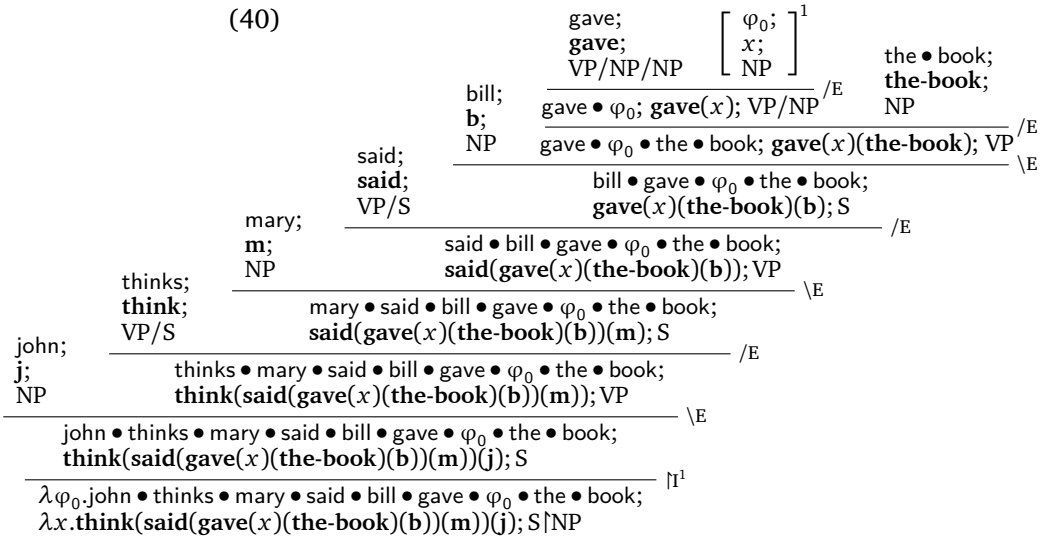
$\mathbf{st} \rightarrow \mathbf{st}$, that is, a function that maps a string to another string). It feeds an empty string ϵ to σ, thereby filling in the embedded gap position, and concatenates the string who with the string thus obtained. For the purpose of exposition, the relevant β-reduction steps are explicitly shown in the dotted line part in (38).

An important property of this analysis is that the gap can be deeply embedded inside the relative clause. Hypothetical reasoning with the vertical slash works exactly in the same way in the simple example above in which the gap corresponds to a local argument position and in the more complex example in (39) in which the gap is located in an embedded clause with multiple levels of embedding.

(39) the guy who John thinks Mary said Bill gave __ the book

The derivation for (39) is shown in (40).

(40)



The addition of a new connective $\uparrow$ necessitates a revision of the definition of syntactic types and the mapping from syntactic to semantic types. In addition, the grammar now recognizes not just simple strings (of type **st**) but also functions that compose such strings in particular ways as admissible prosodic representations of linguistic expressions. We therefore need to define the mapping from syntactic types to prosodic types as well. The new definitions are in (41)–(45).

Syntactic types:

(41) $\mathscr{A} := \{ \text{S, NP, N, ...} \}$ (atomic type)
$\mathscr{D} := \mathscr{A} \mid \mathscr{D}\backslash\mathscr{D} \mid \mathscr{D}/\mathscr{D}$ (directional type)
$\mathscr{T} := \mathscr{D} \mid \mathscr{T}\uparrow\mathscr{T}$ (type)

Semantic types:

(42) (Base Case)
a. $\mathsf{Sem}(\text{NP}) = \mathsf{Sem}(\text{PP}) = e$
b. $\mathsf{Sem}(\text{N}) = e \rightarrow t$
c. $\mathsf{Sem}(\text{S}) = t$

(43) (Recursive Clause)
For any complex syntactic type of the form A/B (or $B\backslash A$, $A\uparrow B$),
$\mathsf{Sem}(A/B)$ $(= \mathsf{Sem}(B\backslash A) = \mathsf{Sem}(A\uparrow B))$ $= \mathsf{Sem}(B) \rightarrow \mathsf{Sem}(A)$

Prosodic types:

(44) (Base Case)
For any directional type $\mathscr{D}$, $\mathsf{Pros}(\mathscr{D}) = \mathsf{st}$ (with st for 'strings').

(45) (Recursive Clause)
For any complex syntactic type $A\uparrow B$ involving $\uparrow$,
$\mathsf{Pros}(A\uparrow B) = \mathsf{Pros}(B) \rightarrow \mathsf{Pros}(A)$.

Note that $\mathscr{D}$ in (41) replaces $\mathscr{T}$ in the earlier definition of syntactic types in (27). The set of syntactic types $\mathscr{T}$ is defined on top of the set of directional types $\mathscr{D}$ (i.e., the complete set of syntactic types in the earlier definition) as in the final clause in (41). This ensures that a vertical slash cannot occur under a directional slash. Thus, S/(S$\uparrow$NP) is not a well-formed syntactic type. One way to make sense of this is to think of it as a 'filter' on uninterpretable prosodic objects. An expression of type X/(Y$\uparrow$Z) would have to concatenate a string to the left of a function of type $\mathsf{st} \rightarrow \mathsf{st}$, but that doesn't make sense.

As the asymmetry between (43) and (45) should make clear, the three slashes /, \ and $\uparrow$ are all functional in the semantic domain, but only $\uparrow$ is functional in the prosodic domain. This asymmetry corresponds to the fact that lambda binding is involved in the prosody only for the Introduction rule for $\uparrow$ (see Section 4.4 for / and \).

Hypothetical reasoning with the directional slashes 4.4

The key notion involved in the analysis of English relative clauses above is hypothetical reasoning, which is essentially a theoretical machinery for 'reasoning about' complex linguistic expressions in which

some material is missing from where it is supposed to appear given the specific lexical specifications of items which make up the complex expressions. In the full version of Hybrid TLG, hypothetical reasoning is generalized to the directional slashes / and \ as well. For the sake of completeness, we show the Introduction rules for / and \, and briefly discuss linguistic applications of these rules.

The Slash Introduction rules for / and \ are formulated as in (46).

(46) a. Forward Slash Introduction

$$\dfrac{\dfrac{\vdots\ [\varphi;\, x;\, A]^n\ \vdots}{\vdots\quad\vdots\quad\vdots}}{b;\, \lambda x.\mathscr{F};\, B/A}\,/\mathrm{I}^n$$

b. Backward Slash Introduction

$$\dfrac{\dfrac{\vdots\ [\varphi;\, x;\, A]^n\ \vdots}{\vdots\quad\vdots\quad\vdots}}{b;\, \lambda x.\mathscr{F};\, A\backslash B}\,\backslash\mathrm{I}^n$$

The difference between the Introduction rule for the vertical slash introduced above in (35a) and these rules is that in (46), the prosodic variable φ for the hypothesis is simply thrown away (instead of being λ-bound). The position of the missing expression is instead recorded in the forward vs. backward slash distinction in the syntactic type.

This is useful when one wants to assign a directional slash type for some string of words in which some material is missing at the periphery, instead of analyzing such expressions with functional prosodic types. For example, for the string *John loves* in the Right-node Raising example in (47), we want to assign the type S/NP so that it is directly conjoinable with another string *Bill hates* of the same type.

(47) [$_{\mathrm{S/NP}}$ John loves], and [$_{\mathrm{S/NP}}$ Bill hates], [$_{\mathrm{NP}}$ Mary].

The derivation for the string *John loves* in type S/NP is shown in (48).

(48)

$$\dfrac{\mathsf{john};\, \mathbf{j};\, \mathrm{NP} \qquad \dfrac{\mathsf{loves};\, \mathbf{love};\, (\mathrm{NP}\backslash\mathrm{S})/\mathrm{NP} \quad [\varphi;\, x;\, \mathrm{NP}]^1}{\mathsf{loves} \bullet \varphi;\, \mathbf{love}(x);\, \mathrm{NP}\backslash\mathrm{S}}\,\backslash\mathrm{E}}{\textcircled{1}\rightarrow\ \dfrac{\mathsf{john} \bullet \mathsf{loves} \bullet \varphi;\, \mathbf{love}(x)(\mathbf{j});\, \mathrm{S}}{\mathsf{john} \bullet \mathsf{loves};\, \lambda x.\mathbf{love}(x)(\mathbf{j});\, \mathrm{S/NP}}\,/\mathrm{I}^1}\,/\mathrm{E}$$

In prose:

- A complete sentence is formed with the hypothetical NP indexed 1. (This much is the same as in the earlier (38).)
- At the next step (①), the hypothesis is withdrawn just as in (38), but here the string variable φ is thrown away, and the derived

type is S/NP (with type **st** prosody). It is this syntactic type that tells us that this is a sentence missing an NP on the right.

Proof term notation of derivations 4.5

To facilitate the ensuing discussion, we introduce here an alternative notation of derivations, one in which a derivation/proof can be written as a single formal object, specifically a lambda term. This corresponds to Abstract Syntax in Abstract Categorial Grammar (de Groote 2001). It exploits the theoretical result in TLG research building on the so-called Curry-Howard Isomorphism (Howard 1969), which states that there is a one-to-one correspondence between proofs and lambda terms in a simply typed lambda calculus. Essentially, an Elimination step (in natural deduction) in a proof corresponds to function application in the lambda calculus and an Introduction step corresponds to lambda abstraction. With Hybrid TLG, this lambda calculus for writing syntactic proofs needs to be extended to distinguish three types of function application ($\text{app}_{/}$, $\text{app}_{\backslash}$, and $\text{app}_{\restriction}$), and three types of lambda abstraction ($\lambda_{/}$, $\lambda_{\backslash}$, and $\lambda_{\restriction}$), corresponding to the three slashes.[11]

As an illustration, consider the derivation (49) (= (38) above) for a simple relative clause from the previous section.

(49)

$$\dfrac{\dfrac{\text{guy};\ \mathbf{guy};\ \text{N} \qquad \dfrac{\lambda\sigma.\text{who} \bullet \sigma(\boldsymbol{\epsilon});\ \lambda P\lambda Q\lambda u.\ Q(u) \wedge P(u);\ (\text{N}\backslash\text{N})\restriction(\text{S}\restriction\text{NP}) \qquad \dfrac{\dfrac{\text{bill};\ \mathbf{b};\ \text{NP} \qquad \dfrac{\text{criticized};\ \mathbf{criticized};\ \text{VP/NP} \qquad \left[\begin{array}{l}\varphi_0;\\ x;\text{NP}\end{array}\right]^1}{\text{criticized} \bullet \varphi_0;\ \mathbf{criticized}(x);\ \text{VP}}\,/\text{E}}{\text{bill} \bullet \text{criticized} \bullet \varphi_0;\ \mathbf{criticized}(x)(\mathbf{b});\ \text{S}}\,\backslash\text{E}}{\lambda\varphi_0.\text{bill} \bullet \text{criticized} \bullet \varphi_0;\ \lambda x.\mathbf{criticized}(x)(\mathbf{b});\ \text{S}\restriction\text{NP}}\,\restriction\text{I}^1}{\text{who} \bullet \text{bill} \bullet \text{criticized} \bullet \boldsymbol{\epsilon};\ \lambda Q\lambda u.Q(u) \wedge \mathbf{criticized}(u)(\mathbf{b});\ \text{N}\backslash\text{N}}\,\restriction\text{E}}{\text{guy} \bullet \text{who} \bullet \text{bill} \bullet \text{criticized} \bullet \boldsymbol{\epsilon};\ \lambda u.\mathbf{guy}(u) \wedge \mathbf{criticized}(u)(\mathbf{b});\ \text{N}}\,\backslash\text{E}$$

We use the same abbreviation of tripartite linguistic signs in the lexicon introduced in Section 2 (= (7)):

[11] This lambda calculus can be thought of as an extension of the bidirectional lambda calculus for the Lambek calculus proposed by Buszkowski (1987) and Wansing (1992). Studying the formal properties of this lambda calculus is an interesting topic on its own, but we leave this task for another occasion.

(50) $\textbf{CRITICIZED}_{\text{TV}}$ = criticized; **past**(**criticize**); TV
$\textbf{WHO}_{(\text{N}\backslash\text{N})\restriction(\text{S}\restriction\text{NP})}$ =
$\lambda\sigma.\text{who} \bullet \sigma(\boldsymbol{\epsilon})$; $\lambda P\lambda Q\lambda u.Q(u) \wedge P(u)$; $(\text{N}\backslash\text{N})\restriction(\text{S}\restriction\text{NP})$
$\textbf{BILL}_{\text{NP}}$ = bill; **b**; NP
$\textbf{GUY}_{\text{N}}$ = guy; **guy**; N

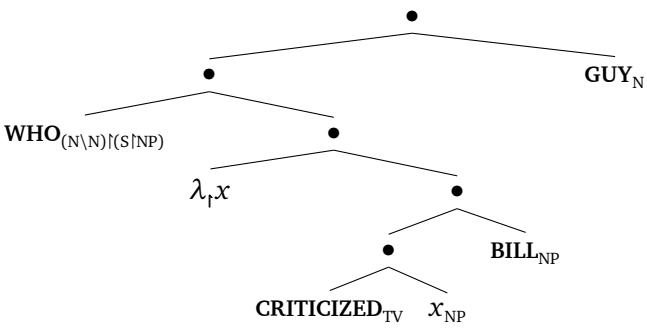
Then, by replacing Slash Elimination by function application and Slash Introduction by lambda abstraction in (49), we obtain the following lambda term, whose syntactic form is isomorphic (i.e., stands in a one-to-one relation) to the natural deduction proof in (49) (the variety of application rule is omitted, since this information is unambiguously recoverable from the syntactic type of the function):

(51) $\textbf{WHO}_{(\text{N}\backslash\text{N})\restriction(\text{S}\restriction\text{NP})}(\lambda_{\restriction}x.\textbf{CRITICIZED}_{\text{TV}}(x_{\text{NP}})(\textbf{BILL}_{\text{NP}}))(\textbf{GUY}_{\text{N}})$

In effect, (51) displays the entire proof narrative exhibited in (49) as a single object: the function corresponding to *criticized* is saturated, with its variable argument undergoing abstraction, yielding an eligible argument for the relative pronoun *who*. Note here that the variable x_{NP} in (51) is a variable in the syntactic logic and is thus formally unrelated to the x in the semantic component of the hypothesis in (49); we use the same variable letter only for expository convenience.

To make it clear that (51) represents underlying semantic composition, and to enhance readability, here is an alternative notation for (51) in the form of a binary tree (already introduced in Section 2):

(52)



Readers familiar with derivational approaches to syntax will recognize a clear resemblance to LF structure. The correspondence to the natural deduction proof tree in (49) should also be easier to see in this format.

The proof term notation is a compact representation of derivations that shows the underlying combinatorics transparently. As we demonstrate below with pied-piping, this is useful in the analyses of complex empirical phenomena involving hypothetical reasoning with the vertical slash (roughly corresponding to 'syntactic movement').

Pied-piping as 'overt and covert' movement 4.6

In the analysis of English relative clauses above, the semantic and syntactic linkage between the extracted material, the relative pronoun and the rest of the sentence is in effect built into the higher-order operator entry for the relative pronoun of type $(N\backslash N)\uparrow(S\uparrow NP)$ in (37). We now consider how this analysis can be extended to pied-piping.

Pied-piping, whimsically named in Ross 1967, 24, is a species of extraction in which a *wh*-pronoun does not directly correspond to a gap within the relative clause but is itself a subconstituent of a larger fronted constituent corresponding to the gap. The following data exemplify the most basic kinds of pied-piping:

(53) a. the guy [to **whom**] John spoke __ yesterday
b. the guy [to **whose** office] John walked __ yesterday
c. the guy [to **whose** sister] John spoke __ yesterday

More elaborate cases can be found, including Ross' example, which makes it clear that the *wh*-word can be embedded arbitrarily deeply.

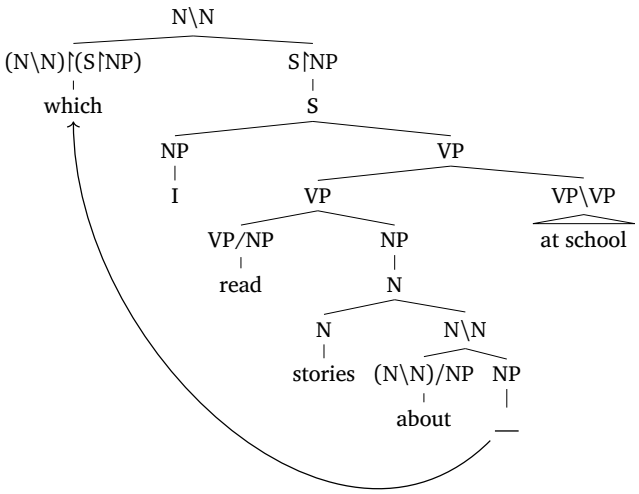
(54) the reports [[the height of the lettering on the covers of **which**] [the government prescribes __]]

Note that the semantic interpretation of pied-piping examples is exactly the same as the corresponding simpler examples in which only the *wh*-word is displaced:

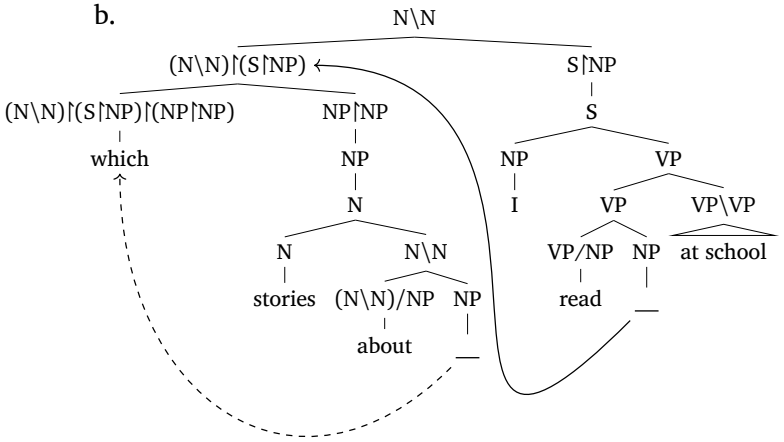
(55) a. Castle Combe is the town [stories about **which**] I read __ at school.
b. Castle Combe is the town **which** I read stories about __ at school.

This correspondence can be graphically represented in the following informal pictures ('overt' and 'covert' movement is represented by solid and dashed lines respectively):

(56) a.



b.



In the case of non-pied-piped relativization (56a), the filler and the gap have the same syntactic type. In contrast, in the pied-piping example (56b), the *wh*-pronoun that triggers relativization is embedded inside the filler, and it is this entire filler phrase that 'binds' the gap in the body of the relative clause. Here, as alluded to by the use of different types of 'movement arrows', the correspondence between the gap and

the filler is a case of 'overt movement', just as with non-pied-piped relativization. By contrast, the identification of the whole *wh*-phrase that contains the *wh*-word as the 'operator' that triggers relativization is mediated by a 'covert movement'-like operation. In the latter, the string of the *wh*-word is embedded inside the filler phrase.

This can be formalized precisely by modifying the lexical entry for the *wh*-operator as in (57) (the key idea here is due to Morrill (1994)).

(57) $\lambda\sigma_1\lambda\sigma_2.\sigma_1(\mathsf{whom}) \bullet \sigma_2(\epsilon)$;
$\lambda F\lambda P\lambda Q\lambda x.P(F(x)) \wedge Q(x)$; $(\mathrm{N}\backslash\mathrm{N})\uparrow(\mathrm{S}\uparrow\mathrm{X})\uparrow(\mathrm{X}\uparrow\mathrm{NP})$

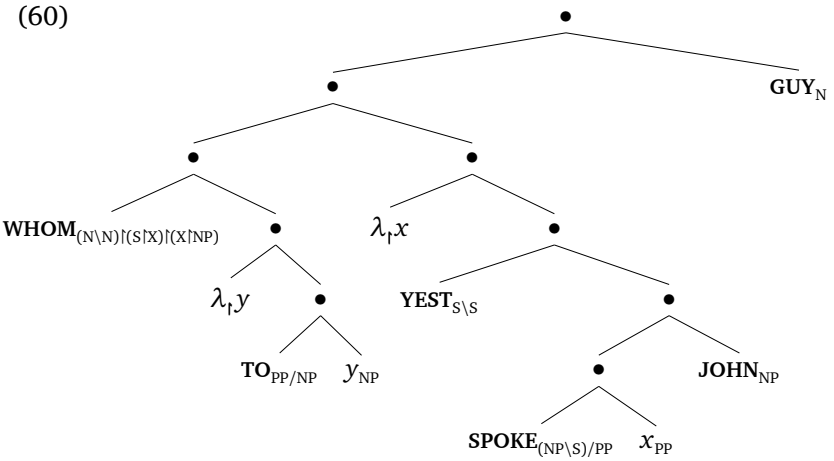
This says that the relative pronoun takes two arguments, some expression of type X missing an NP and an S missing an X, and then becomes a nominal modifier. A sample derivation for (53a) using this entry is shown in (58) (in natural deduction) and (59)/(60) (in the proof term format). Here, since the fronted phrase is a PP, X is instantiated as PP.

(58)

$$\dfrac{\dfrac{\lambda\sigma_1\lambda\sigma_2.\sigma_1(\mathsf{whom}) \bullet \sigma_2(\epsilon);\ \lambda F\lambda P\lambda Q\lambda x.P(F(x)) \wedge Q(x);\ (\mathrm{N}\backslash\mathrm{N})\uparrow(\mathrm{S}\uparrow\mathrm{X})\uparrow(\mathrm{X}\uparrow\mathrm{NP}) \qquad \dfrac{\dfrac{\mathsf{to};\ \lambda x.x;\ \mathrm{PP}/\mathrm{NP} \qquad \left[\begin{array}{l}\varphi_2;\\ y;\mathrm{NP}\end{array}\right]^2}{\mathsf{to} \bullet \varphi_2;\ y;\ \mathrm{PP}}\,/\mathrm{E}}{\lambda\varphi_2.\mathsf{to} \bullet \varphi_2;\ \lambda y.y;\ \mathrm{PP}\uparrow\mathrm{NP}}\,\uparrow\mathrm{I}^2}{\lambda\sigma_2.\mathsf{to} \bullet \mathsf{whom} \bullet \sigma_2(\epsilon);\ \lambda P\lambda Q\lambda x.P(x) \wedge Q(x);\ (\mathrm{N}\backslash\mathrm{N})\uparrow(\mathrm{S}\uparrow\mathrm{PP})}\,\uparrow\mathrm{E} \qquad \dfrac{\left[\begin{array}{l}\varphi_1;\\ x;\mathrm{PP}\end{array}\right]^1}{\begin{array}{c}\vdots\\ \lambda\varphi_1.\mathsf{john} \bullet \mathsf{spoke} \bullet \varphi_1 \bullet \mathsf{yesterday};\\ \lambda x.\mathbf{yest}(\mathbf{spoke}(x)(\mathbf{j}));\\ \mathrm{S}\uparrow\mathrm{PP}\end{array}}\,\uparrow\mathrm{I}^1}{\mathsf{to} \bullet \mathsf{whom} \bullet \mathsf{john} \bullet \mathsf{spoke} \bullet \mathsf{yesterday};\ \lambda Q\lambda x.\mathbf{yest}(\mathbf{spoke}(x)(\mathbf{j})) \wedge Q(x);\ \mathrm{N}\backslash\mathrm{N}}\,\uparrow\mathrm{E}$$

(59) $\mathbf{WHOM}_{(\mathrm{N}\backslash\mathrm{N})\uparrow(\mathrm{S}\uparrow\mathrm{X})\uparrow(\mathrm{X}\uparrow\mathrm{NP})}$
$(\lambda_\uparrow y.\mathbf{TO}_{\mathrm{PP}/\mathrm{NP}}(y_{\mathrm{NP}}))(\lambda_\uparrow x.\mathbf{YEST}_{\mathrm{S}\backslash\mathrm{S}}(\mathbf{SPOKE}_{(\mathrm{NP}\backslash\mathrm{S})/\mathrm{PP}}(x_{\mathrm{PP}})(\mathbf{JOHN}_{\mathrm{NP}})))$

(60)



Note that this analysis involves two instances of hypothetical reasoning, corresponding to the 'overt' and 'covert' movement operations in the informal diagram in (56b). The hypothetical reasoning with the PP (indexed 1 in (58) and x_{PP} in (59)/(60)) is for forming a gapped sentence of type S⇂PP that serves as the body of the relative clause. The hypothetical reasoning involving the NP hypothesis (indexed 2 in (58) and y_{NP} in (59)/(60)) is for identifying the location of the relative pronoun inside the fronted constituent *to whom*. The relativization operator defined in (57) fills in an empty string and the string of the relative pronoun (i.e., the string *whom*) in the positions of the two lambda-bound variables φ_1 and φ_2, reflecting the 'overt' and 'covert' movement statuses of the two hypothetical reasoning steps involved. In Hybrid TLG, 'covert' and 'overt' movement are handled by the same formal mechanism, and the difference between the two merely consists in whether an overt string is substituted for the bound variable position in the prosodic function that is given as an argument to the higher-order operator.

Since the 'in-situ' operator relationship between the relative pronoun and the fronted expression containing it is mediated by ⇂, we predict that the *wh*-pronoun can be embedded inside the fronted constituent arbitrarily deeply. Thus, Ross's (1967) example can be accounted for in the same way as the simpler PP pied-piping example in (58) above. We show the derivation in proof term notation:

(61) $\mathbf{WHICH}_{(\mathrm{N}\backslash\mathrm{N})\restriction(\mathrm{S}\restriction\mathrm{X})\restriction(\mathrm{X}\restriction\mathrm{NP})}$
$(\lambda_{\restriction} y.\mathbf{THE}_{\mathrm{NP/N}}(\mathbf{HEIGHT}_{\mathrm{N/PP}}(\mathbf{OF}_{\mathrm{PP/NP}}$
$(\mathbf{THE}_{\mathrm{NP/N}}(\mathbf{ON}_{(\mathrm{N}\backslash\mathrm{N})/\mathrm{NP}}(\mathbf{THE}_{\mathrm{NP/N}}(\mathbf{COVERS}_{\mathrm{N/PP}}(\mathbf{OF}_{\mathrm{PP/NP}}(y_{\mathrm{NP}}))))$
$(\mathbf{LETTERING}_{\mathrm{N}})))))$
$(\lambda_{\restriction} x.\mathbf{PRESCRIBES}_{(\mathrm{NP}\backslash\mathrm{S})/\mathrm{NP}}(x_{\mathrm{NP}})(\mathbf{THE}_{\mathrm{NP/N}}(\mathbf{GVT}_{\mathrm{N}})))$
$=$ the • height • of • the • lettering • on • the • covers•
of • which • the • government • prescribes;
$\lambda P\lambda x.P(x) \wedge \mathbf{prescribe}(\mathbf{the}(\mathbf{height}(\mathbf{the}(\mathbf{on}(\mathbf{the}$
$(\mathbf{covers}(x)))(\mathbf{lettering})))))(\mathbf{the}(\mathbf{gvt}));\mathrm{N}\backslash\mathrm{N}$

Here, X is instantiated as NP. The question of which syntactic type can be pied-piped is a rather thorny issue. As noted by Arnold and Godard (2021), even a descriptively correct generalization for a well-

studied language like English is unclear. We won't attempt to address this issue, since the analysis of pied-piping itself is not our central goal.

5 EXTRACTION PATHWAY MARKING AS PROOF STRUCTURE MARKING

Having reviewed the system of Hybrid TLG, we are now ready to present the full formal analysis of EPM. We start our illustration with the Iringlish case in Section 5.1 (which is mostly a review of the proposal already presented in Section 3.2). This is followed by an illustration of a wider range of options that other languages exploit for the purpose of EPM encoding (Section 5.2). Here, we focus in particular on the floating quantifier *all* in Irish English and information structure-sensitive word-order encoding in Dinka, while touching on various related strategies displayed by other languages along the way. This discussion is meant to demonstrate that our proof-theoretic reconceptualization of the notion of cyclicity has a broad empirical coverage with some interesting semantically-oriented typological implications (discussed briefly in Section 6). Section 5.3 then briefly considers implications for other phenomena pertaining to cyclicity such as reconstruction effects. The final part of this section (Section 5.4) offers a brief comparison with an approach to EPM in HPSG, which dispenses with cyclic movement but encodes the effect by feature propagation. We believe that the discussions in this section will clarify further the ways in which our approach inherits the key ideas of the earlier accounts as well as ways in which it can be seen to offer new insights.

5.1 *Accounting for extraction pathway marking*

Since we have already presented the analysis of Iringlish in informal terms in Section 3, here, for the most part we just reproduce the formal lambda terms corresponding to the informal tree diagrams in Section 3.3. This is followed by some additional discussions of residual issues (on Iringlish and other languages).

The proof term notation for the tree in (20) (for (62)) can be written as (63).

(62) the man **aL** [I said **aL** [I thought **aL** [__ would be there]]]

(63) $\mathbf{REL}_{(N\backslash N)\uparrow(S'\uparrow NP_{+wh})}$
$(\lambda_{\uparrow} x.\mathbf{AL}_{S'/S}(\mathbf{SAID}_{(NP\backslash S)/S'}$
$(\mathbf{AL}_{S'/S}(\mathbf{THOUGHT}_{(NP\backslash S)/S'}$
$(\mathbf{AL}_{S'/S}(\mathbf{WBT}_{NP\backslash S}(x_{NP_{+wh}})))(\mathbf{I}_{NP})))(\mathbf{I}_{NP})))$

Here, each token of *aL* applies to a clausal complement containing a free NP_{+wh} variable, and hence is legal.

The bad cases in (22) can be reproduced in the form of proof terms as in (64).

(64) a. $\mathbf{REL}_{(N\backslash N)\uparrow(S'\uparrow NP_{+wh})}$
$(\lambda_{\uparrow} x.\mathbf{AL}_{S'/S}(\mathbf{SAID}_{(NP\backslash S)/S'}$
$(\mathbf{AL}_{S'/S}(\mathbf{THOUGHT}_{(NP\backslash S)/S'}$
$(\mathbf{GON}_{S'/S}(\mathbf{WBT}_{NP\backslash S}(x_{NP_{+wh}})))(\mathbf{I}_{NP})))(\mathbf{I}_{NP})))$

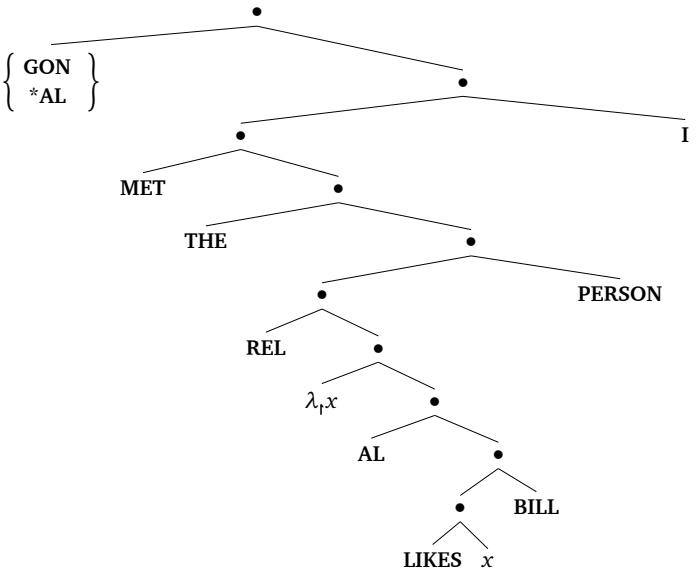
b. $\mathbf{REL}_{(N\backslash N)\uparrow(S'\uparrow NP_{+wh})}$
$(\lambda_{\uparrow} x.\mathbf{GON}_{S'/S}(\mathbf{SAID}_{(NP\backslash S)/S'}$
$(\mathbf{AL}_{S'/S}(\mathbf{THOUGHT}_{(NP\backslash S)/S'}$
$(\mathbf{AL}_{S'/S}(\mathbf{WBT}_{NP\backslash S}(x_{NP_{+wh}})))(\mathbf{I}_{NP})))(\mathbf{I}_{NP})))$

A further prediction of this approach is that when extraction terminates in an embedded clause, the complementizer in a higher structure will be *goN*, rather than *aL*. We illustrate this point with the following (artificial) example:

(65) I said $\left\{ \begin{array}{c} \mathbf{goN} \\ \mathbf{*aL} \end{array} \right\}$ I met the person$_i$ **aL** [Bill likes $\_\_{}_i$].

In (66), the variable x corresponding to the trace in the embedded relative clause is bound by the lambda operator in the subterm given as an argument to the relativization operator. Thus, the proof term given as an argument to the topmost *aL/goN* contains no free variable. Hence, only *goN* is allowed in the higher clause.

(66)



A case we did not discuss explicitly in Section 3.2 is adjunct extraction. This is completely parallel to extraction of arguments.[12] In an example such as (67), the extracted adjunct semantically modifies the embedded clause. Thus, a hypothetical clausal modifier of type $(S\backslash S)_{+wh}$ is posited in the lower clause as in (68)/(69).[13]

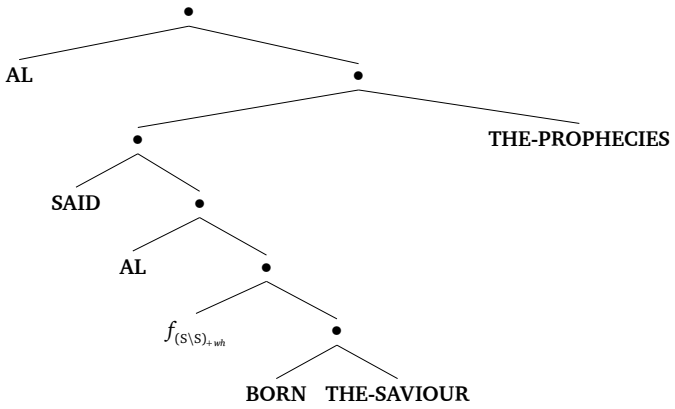
(67) It was in Bethlehem **aL** [the prophecies said **aL** [the Saviour would be born __]]

(68) $\mathbf{AL}_{S'/S}(\mathbf{SAID}_{(NP\backslash S)/S'}$
$(\mathbf{AL}_{S'/S}(f_{(S\backslash S)_{+wh}}(\mathbf{BORN}_{NP\backslash S}(\mathbf{THE\text{-}SAVIOUR}_{NP}))))(\mathbf{THE\text{-}PROPHECIES}_{NP}))$

[12] Adjunct extraction poses an interesting theoretical issue in lexicalist theories of syntax such as HPSG and (some variants of) CG (see, e.g., Hukari and Levine 1995), since in such theories, there is an asymmetry between arguments and adjuncts in that the former is an argument of a lexical verb but the latter is standardly a function that takes a verbal projection as an argument. Thus, the pattern in (62) presents a non-trivial issue for a feature-percolation analysis sensitive to valence information of the sort briefly discussed in Section 5.3 below (see Bouma *et al.* 2001 and Levine and Hukari 2006 for details).

[13] Admitting the syntactic type $(S\backslash S)_{+wh}$ necessitates a move in the underlying theory in which not just atomic types but also complex types can be specified for (at least certain) syntactic features. This may involve some major reworking of the feature system in TLG, but we leave this task for future work.

(69)



Here again, until the variable f (of type $(S\backslash S)_{+wh}$) is bound, the right form of the complementizer is *aL*, so it is correctly predicted that the two occurrences of *aL* in (67) cannot be replaced by *goN*.

The analysis of complementizer marking in Irish presented above exploits the fact that 'movement' phenomena are analyzed by hypothetical reasoning in TLG and that unwithdrawn hypotheses can be formally treated as unbound variables in the lambda calculus representing proofs. The same approach can be directly extended to cases in which EPM is registered by phenomena that affect the 'clause structure', such as the inversion strategy in Belfast English (and perhaps in French, too, but see the critique of Kayne and Pollock 1978 by Bonami *et al.* (1999)).

(70) What did John say [$_{CP}$ __ **did** Mary claim [$_{CP}$ __ **had** John feared __]]?

Assuming that Henry's (1995) characterization of the empirical facts is correct, Belfast English registers extraction pathways by subject-auxiliary inversion consistently.

In lexicalist theories of syntax such as categorial grammar, the standard analysis of inversion involves lexical encoding of the inverted order in the syntactic type of the auxiliary verb (Gazdar *et al.* 1982; Sag *et al.* 2020; Kubota and Levine to appear). For example, in addition to the uninverted, normal word-order variant in (71a) (in which an auxiliary essentially takes a nonfinite VP and returns a finite VP), we have a lexically related alternative entry in (71b) in which it combines with the subject first before combining with its nonfinite VP complement.

(71) a. had; $\lambda F.F$; $(NP\backslash S_{fin})/(NP\backslash S_{bse})$
b. had; $\lambda F.F$; $S_{inv}/(NP\backslash S_{bse})/NP$

The registering of EPM via inversion is straightforward in this type of lexicalist analysis of auxiliaries. In Belfast English, the auxiliary verb entries of the sort in (71) come with additional restrictions that reference the existence of free variables in their $NP\backslash S_{bse}$ syntactic arguments, just like the two complementizer forms in Irish in (18).

5.2 *Extraction pathway marking in other languages*

Having provided an analysis of the basic patterns of EPM, we now turn to the question of whether this analysis is fully general. For this purpose, we critically examine the recent claim by van Urk and Richards (2015) and van Urk (2020) that *both* successively cyclic movement and feature percolation are needed to capture the entire patterns of EPM. According to van Urk and Richards (2015), the crucial piece of evidence comes from the patterns displayed by Dinka. The apparent violation of the V2 word order in the language exceptionally observed at *wh*-extraction pathways provides evidence for actual movement of the *wh*-phrase. However, the 'long-distance' plural agreement cannot be accounted for by movement alone, and requires a feature checking (or feature percolation) mechanism of some sort. Van Urk (2020) summarizes facts from a wider range of languages for each type of evidence.

To state the conclusion first, while we agree with these authors that these phenomena call for some mechanism in the grammar for keeping track of the identity of the gap before the filler-gap linkage is established, the relevant facts can be analyzed adequately by what we have already proposed, together with independently motivated properties of the specific morpho-syntactic phenomena that exhibit EPM effects. Among the two types of alleged evidence for distinct mechanisms, the 'feature checking' evidence can be dealt with by a slight extension of the analysis of the Irish complementizer marking pattern. We briefly demonstrate this point in Section 5.2.1. After that, we turn to the main task in this section, focusing on two types of 'movement evidence' reported in van Urk and Richards 2015 and van Urk 2020,

specifically, Dinka word order (Section 5.2.2) and Irish English floating quantifier *all* (Section 5.2.3). Importantly, a key component of van Urk and Richards's (2015) claim is that Dinka exhibits the 'feature percolation' pattern and the 'movement' pattern within a single language. We counter this claim by showing that the two patterns found in this language (plural marking and word order) can be dealt with by making different lexical items in the language sensitive to essentially the same type of information.

5.2.1 A note on 'agreement' type extraction pathway marking effects

Cases of EPM in which the marking is sensitive to some particular syntactic or semantic feature of the extracted expression, such as the plural marking morphology in Dinka reported in van Urk and Richards 2015, perhaps requires some discussion, before we tackle the main issue of the movement-type evidence for EPM. Here, we show that such cases can be analyzed essentially by the same approach we proposed for Irish complementizer marking, together with the feature-based account of agreement standardly assumed in lexicalist syntax (including TLG).

For the purpose of illustration, suppose that Iringlish had morphological indication of the plurality of the extracted item realized as reduplicative morphology in the form of an intermediate verb. Agreement is handled via features encoded in syntactic categories in lexicalist theories of syntax. Using this feature-based analysis of agreement, a plural-gap variant of the verb *think* can be defined as follows:

(72) **THOUGHT-PL**$_{\mathrm{VP}/\mathrm{S}'}$ = thought-thought; **thought**; VP/S'
where for any α, **THOUGHT-PL**(α) is defined only if $fv_{\mathrm{X}_{+wh}}(\alpha) \neq \emptyset$ and the singleton element of $fv_{\mathrm{X}_{+wh}}(\alpha)$ has type NP_{+pl}

(73) the { a.*man / b. men } aL [I **thought-thought** aL [__ would be there]]

Since the gap NP and the head noun are required to agree in number by the relativization operator, in (73a) the gap NP has type NP_{-pl} and in (73b) it has type NP_{+pl}, yielding the subterms in (74a) and

(74b), respectively, as arguments to (72). Only the former satisfies the definedness condition for (72), correctly capturing the pattern in (73).

(74) a. $\mathbf{AL}_{S'/S}(\mathbf{WBT}_{NP\backslash S}(x_{NP_{-pl}}))$
b. $\mathbf{AL}_{S'/S}(\mathbf{WBT}_{NP\backslash S}(x_{NP_{+pl}}))$

V2 word order in Dinka 5.2.2

Van Urk and Richards (2015) present the following pattern of extraction pathway marking reflected in V2 word order in Dinka as evidence for an actual movement of a copy of the *wh*-phrase in successive cyclicity. We reproduce the relevant pattern in Dinklish, another hypothetical dialect of English which mimicks (the relevant part of) Dinka syntax with an English lexicon.

First, (75) shows that normally embedded clauses exhibit the V2 word order, and that leaving the preverbal position empty is not allowed.

(75) a. Bill$_j$ thinks __$_j$ ke [Mary$_i$ bought __$_i$ the book].
'Bill thinks that Mary bought the book.'
b. *Bill$_j$ thinks __$_j$ ke [__ bought Mary the book].

But there is a systematic exception to this V2 word order requirement. The preverbal position can, and in fact must, be empty when it is crossed by a *wh*-dependency chain. This is demonstrated by (76).

(76) a. Who$_i$ thought John ke [__$_i$ said Mary ke [__$_i$ criticized Bill __$_i$]]?
'Who did John think Mary said Bill criticized __?'
b. *Who$_i$ thought John ke [**Mary**$_j$ said __$_j$ ke [__$_i$ criticized Bill __$_i$]]?
c. *Who$_i$ thought John ke [__$_i$ said Mary ke [**Bill**$_j$ criticized __$_j$ __$_i$]]?
d. *Who$_i$ thought John ke [**Mary**$_k$ said __$_k$ ke [**Bill**$_j$ criticized __$_j$ __$_i$]]?

(76a) is grammatical since the preverbal positions in the most embedded and intermediate clauses are both left unoccupied. By contrast, in the ungrammatical (76b–d), either the preverbal position in the lowest or the intermediate clause (or both) is occupied by an overt NP.

Van Urk and Richards (2015) characterize the preverbal position as Spec,CP. According to them, the pattern in (76) falls out immediately if Spec,CP is an intermediate landing site of the moved *wh*-phrase.

However, there is an alternative account of this distributional pattern that doesn't rely on actual movement of a *wh*-phrase, in which the semantic effect of extraction is taken to be a key component of the explanation. The key idea is that the preverbal position in Dinklish (or Dinka) corresponds to the 'variable' slot in the abstract predicate-argument structure underlying the topic/comment structure in ordinary sentences and the focus/background structure in *wh*-questions. To make this idea more concrete, we make the following assumptions:

(77) a. Every clause must be associated with at most one 'most prominent' element.

b. The preverbal position is the designated position for the prominent element, and is licensed through $\uparrow$.

c. As a consequence of (77a,b) when $\uparrow$-Introduction applies to produce a predicate-argument structure underlying V2 syntax, there has to be exactly one unwithdrawn hypothesis (corresponding to the element carrying prominence).

To see how this works, consider first the following simple 'Dinklish' sentence with local topicalization:

(78) Bill$_i$ gave __$_i$ Mary the book.

(79)

$$\dfrac{\dfrac{\lambda\varphi_1\lambda\sigma.\ \varphi_1 \bullet \sigma(\boldsymbol{\epsilon});\ \lambda z\lambda R.R(z);\ S{\uparrow}(S{\uparrow}NP){\uparrow}NP \qquad \textsf{bill};\ \mathbf{b};\ NP}{\lambda\sigma.\textsf{bill} \bullet \sigma(\boldsymbol{\epsilon});\ \lambda R.R(\mathbf{b});\ S{\uparrow}NP}\ {\uparrow}E \qquad \dfrac{\dfrac{\dfrac{\textsf{gave};\ \lambda y\lambda x\lambda w.\ \mathbf{gave}(x)(w)(y);\ S/NP/NP/NP \qquad \begin{bmatrix}\varphi_1;\\ v;\\ NP\end{bmatrix}^1}{\textsf{gave} \bullet \varphi_1;\ \lambda x\lambda w.\mathbf{gave}(x)(w)(v);\ S/NP/NP \qquad \textsf{mary};\ \mathbf{m};\ NP}\ /E \qquad \vdots}{\textsf{gave} \bullet \varphi_1 \bullet \textsf{mary};\ \lambda x.\mathbf{gave}(\mathbf{m})(w)(v);\ S/NP \qquad \textsf{the} \bullet \textsf{book};\ \iota(\mathbf{book});\ NP}\ /E}{\textsf{gave} \bullet \varphi_1 \bullet \textsf{mary} \bullet \textsf{the} \bullet \textsf{book};\ \mathbf{gave}(\mathbf{m})(\iota(\mathbf{book}))(v);\ S}\ {\uparrow}I^1}{\lambda\varphi_1.\textsf{gave} \bullet \varphi \bullet \textsf{mary} \bullet \textsf{the} \bullet \textsf{book};\ \lambda v.\mathbf{gave}(\mathbf{m})(\iota(\mathbf{book}))(v);\ S{\uparrow}NP}\ {\uparrow}E$$

$$\textsf{bill} \bullet \textsf{gave} \bullet \varphi \bullet \textsf{mary} \bullet \textsf{the} \bullet \textsf{book};\ \mathbf{gave}(\mathbf{m})(\iota(\mathbf{book}))(\mathbf{b});\ S$$

(80) $\mathbf{TOP}_{S{\uparrow}(S{\uparrow}NP){\uparrow}NP}(\mathbf{BILL}_{NP})(\lambda_{\uparrow}x.\mathbf{GAVE}_{S/NP/NP/NP}(x_{NP})(\mathbf{MARY}_{NP})(\mathbf{THE\text{-}BOOK}_{NP}))$

At the step ↑-Introduction applies, there is exactly one free variable x (corresponding to the unwithdrawn hypothesis indexed 1), so the derivation succeeds. Since this hypothesis corresponds to the subject argument of the verb *gave*, we get a subject topicalization sentence.

Consider next the following minimal pair (= (75)), which shows that an embedded topic position cannot remain empty:

(81) a. Bill thinks __ ke [Mary bought __ the book].
'Bill thinks that Mary bought the book.'
b. *Bill thinks __ ke [__ bought Mary the book].
'Bill thinks that Mary bought the book.'

To account for this pattern (and also the *wh*-dependency patterns below), we assume that the complementizer *ke* has the role of ensuring the condition (77a) above, which can be made explicit as in (82).

(82) *Ke* imposes the restriction that there is exactly one free variable in its complement.

As we show immediately below, in the normal topicalization example, after *ke* checks the existence of a free variable, the variable gets bound by ↑-Introduction as usual, and the result is then fed to the topicalization operator; otherwise, that is, when there is a filler corresponding to an embedded gap in a higher clause, *ke* simply passes the free variable upstairs.

For (81), what goes wrong in (81b) is that at the point *ke* combines with the embedded clause, both of the argument positions are occupied by full NPs as in (83a). This violates the condition on *ke* in (82), hence the derivation fails. By contrast, in the case of the topicalization example (81a), the underlined subproof in (83b) satisfies (82), with the free variable x_{NP} which then gets bound by the topicalization operator that licenses the overt NP *Mary* in the clause initial position.

(83) a. $\mathbf{KE}_{\mathrm{S'/S}}(\mathbf{BOUGHT}_{\mathrm{S/NP/NP}}(\mathbf{THE\text{-}BOOK}_{\mathrm{NP}})(\mathbf{MARY}_{\mathrm{NP}}))$
b. $\mathbf{TOP}_{\mathrm{S}\uparrow(\mathrm{S}\uparrow\mathrm{NP})\uparrow\mathrm{NP}}$
$(\lambda_{\uparrow} x.\mathbf{KE}_{\mathrm{S'/S}}\underline{(\mathbf{BOUGHT}_{\mathrm{S/NP/NP}}(\mathbf{THE\text{-}BOOK}_{\mathrm{NP}})(x_{\mathrm{NP}}))})(\mathbf{MARY}_{\mathrm{NP}})$

Assuming that the same constraint is operative in more complex sentences involving long-distance extraction of a *wh*-phrase, the pattern in (76) falls out from the assumptions already made. As noted

above, all the preverbal positions in intermediate clauses crossed by filler-gap linkage have to be empty:

(84) a. I wonder who$_i$ thought John ke [__$_i$ said Mary ke [__$_i$ criticized Bill __$_i$]].
'I wonder who John thought Mary said Bill criticized __.'

b. *I wonder who$_i$ thought John ke [Mary$_j$ said __$_j$ ke [__$_i$ criticized Bill __$_i$]].
'I wonder who John thought Mary said Bill criticized __.'

We start with the analysis of the grammatical example (84a). Note first that the subproof for the most deeply embedded clause satisfies both (77) and (82), since it contains exactly one hypothesis x_{NP}.

(85) $\mathbf{KE}_{\mathrm{S'/S}}(\underline{\mathbf{CRITICIZED}_{\mathrm{S/NP/NP}}(x_{\mathrm{NP}})(\mathbf{BILL}_{\mathrm{NP}})})$

The same process is repeated in the upstairs clause, yielding (86), again satisfying the relevant conditions at the intermediate clause headed by *said*:

(86) $\mathbf{KE}_{\mathrm{S'/S}}(\mathbf{SAID}_{\mathrm{S/NP/S'}}(\mathbf{KE}_{\mathrm{S'/S}}(\mathbf{CRITICIZED}_{\mathrm{S/NP/NP}}(x_{\mathrm{NP}})(\mathbf{BILL}_{\mathrm{NP}})))(\mathbf{MARY}_{\mathrm{NP}}))$

Finally, at the matrix level, the hypothesis is withdrawn to yield S⇂NP, which is then given as an argument to the *wh*-operator:

(87) $\mathbf{WHO}_{\mathrm{Q}\restriction(\mathrm{S}\restriction\mathrm{NP})}(\lambda_{\restriction}x.\mathbf{THOUGHT}_{\mathrm{S/S'/NP}}(\mathbf{JOHN}_{\mathrm{NP}})(\mathbf{KE}_{\mathrm{S'/S}}(\mathbf{SAID}_{\mathrm{S/NP/S'}}(\mathbf{KE}_{\mathrm{S'/S}}(\mathbf{CRITICIZED}_{\mathrm{S/NP/NP}}(x_{\mathrm{NP}})(\mathbf{BILL}_{\mathrm{NP}}))(\mathbf{MARY}_{\mathrm{NP}})))))$

Turning now to the ungrammatical (84b), the offending structure is the subproof for the intermediate clause headed by *said*, where the preverbal position is occupied by the local subject *Mary* of that clause, instead of being left empty. As in the above (81a) (with derivation in (83b)), in order to license an overt NP in the topic position, we need to do hypothetical reasoning as in (88). But the underlined part violates the condition on *ke* in (82), since this subproof has two variables x_{NP} (corresponding to the *wh*-filler) and y_{NP} (for the local topic).

(88) $\mathbf{TOP}_{\mathrm{S}\restriction(\mathrm{S}\restriction\mathrm{NP})\restriction\mathrm{NP}}(\lambda_{\restriction}y.\mathbf{KE}_{\mathrm{S'/S}}(\underline{\mathbf{SAID}_{\mathrm{S/NP/S'}}(\mathbf{KE}_{\mathrm{S'/S}}(\mathbf{CRITICIZED}_{\mathrm{S/NP/NP}}(x_{\mathrm{NP}})(\mathbf{BILL}_{\mathrm{NP}})))(y_{\mathrm{NP}})}))$

To summarize, the Dinka V2 word order pattern in (76) (in Dinklish) can be explained by an interaction of the topicalization operator and *wh*-extraction. The ungrammatical cases all violate the constraint that there has to be exactly one 'prominent' element in a clause. Since both topicalization and *wh*-extraction exploit hypothetical reasoning at the syntax-semantics interface to identify a particular expression as the 'prominent' element with respect to the respective constructions (where 'prominent' corresponds to focus in *wh*-extraction and topic in topicalization), we predict the same pattern as van Urk and Richards (2015), without treating the preverbal position as a particular type of syntactic projection targeted by cyclic movement.

Linking the interpretation of a variable to discourse prominence may seem like a stipulative association of a syntactic restriction on semantic interpretation with an information-structural property of a dynamic pragmatic background. But increasingly, it is becoming evident that such associations must be recognized, in the interest of empirical generality. For example, this is precisely the kind of condition that Toosarvandani (2016) identifies as the basis for configurational restrictions on the distribution of Gapping in English. In still more recent work, Barros and Frank (2023) have shown that apparently purely syntactic restrictions on the interpetation of multiple sluicing (for which a phase-based analysis was attempted in an earlier work by Grano and Lasnik (2018)) are best understood in terms of discourse prominence status holding between discourse referents in material separated by a clause boundary. Note in particular here that there is a quite suggestive parallel with our proposal for Dinka: in both analyses, there is a prominence relationship established in higher clauses which determines how a variable – corresponding to a bound pronoun in the English data and a reserved preverbal position in Dinka – can be interpreted. We take this sort of dependency relationship to point to a principled basis for the condition in (82).

Irish English *all* 5.2.3

McCloskey (2000) argues that the Ulster subdialect of Irish English allows the extracted operator *what all* to jettison the quantifier-like *all* at various points along a Spec-to-Spec series of local extraction steps, giving tangible evidence that the extracted *wh*-phrase has passed

through those steps to arrive at its final landing site. His evidence for this analysis includes the set of data in (89)–(91).

(89) a. What **all** did you get __ for Christmas?
b. Who **all** did you meet __ when you were in Derry?

(90) a. What did you get **all** __ for Christmas?
b. Who did you meet **all** __ when you were in Derry?

(91) a. What **all** did he say (that) he wanted __?
b. What did he say (that) he wanted __ **all**?
c. What did he say **all** (that) he wanted __?

On McCloskey's reasoning, the semantic identity of the floating and non-floating variants of *what/who all* sentences in (89) vs. (90) justifies an analysis in which *what/who all* is 'underlyingly' a unit. On the other hand, as illustrated in (91), the apparently free-floating *all* appears at exactly the points that correspond either to the *wh*-element's site of origin (as in (91b)) or to an intermediate Spec,CP position on the extraction pathway (as in (91c)). McCloskey then takes the distribution of *all* as (at least indirect) evidence for cyclic movement.

In what follows, we sketch an alternative explanation of these facts which essentially takes *all* to be an adverb, building on Sag and Levine (2006), who offer an argument involving the parallel between Irish English *all* and *exactly/precisely* in Standard American English. We refine the connection between the adverbial syntax of *all* and the semantic effect that it imposes on the interpretation of the fronted *wh*-word, an aspect that remains somewhat vague in the Sag/Levine account. We take *all* to be syntactically a VP adverb which imposes a certain semantic restriction on a free variable in its argument. This latter semantic effect is what gives rise to the apparent synonymy between the floating and non-floating variants of *what/who … all*. Here again, our account crucially makes reference to the intermediate status of the proof, in such a way that the semantic interpretation of the free variable (unwithdrawn hypothesis) plays a key role.

One piece of evidence for the assumption that stranded *all* is an adverb comes from data such as the following:

(92) ?What did you put in the drawer __ **all** (yesterday)?

On the VP modifier analysis, the position of *all* in (92) is naturally expected. By contrast, on McCloskey's (2000) movement-based analysis, (92) has to be analyzed as first involving a local movement of *what all* to the post-PP position (which is prohibited for overt, non-*wh*-NPs). However, such an analysis seems implausible given the lack of any independent evidence for the supposed movement operation.[14]

---

[14] Further support for the VP adverb analysis of *all* comes from the distributional parallel between the non-remnant adverb *precisely* and the floating *all*. Note first that *precisely* appears to have a very similar distribution as *all*, occuring in both the post-*wh* position and the 'in-situ' position:

(i) a. What **precisely** do you want __?
b. What do you want __ **precisely**?

However, as McCloskey himself notes, a closer inspection makes it clear that *precisely* cannot plausibly be analyzed as a *wh*-remnant:

(ii) a. *What did he say yesterday **precisely** that he wanted? [on the same reading as (iib)]
b. What **precisely** did he say yesterday that he wanted?

If *precisely* were a *wh*-remnant on a par with *all*, then (iia) should have a reading equivalent to (iib), with *precisely* being stranded at an intermediate landing site. However, (iia) clearly lacks such a reading.

Yet despite this clear difference in the *wh*-remnant status, *precisely* and *all* share a remarkable similarity in terms of their syntactic distribution as VP-internal adverbs, as shown by the following examples:

(iii) a. *What did he say {**precisely**/**all**} to {him/his students} that he wanted to buy __?
b. ?What did he say to {him/his students} {**precisely**/**all**} that he wanted to buy __?

This distributional parallel between *precisely* and *all* indicates that the pre-complementizer distribution of *all* that McCloskey takes as sufficient evidence for the Spec,CP remnant status of *all* can be accounted for equally naturally by simply assuming that it is syntactically a VP adverb that obeys the same word-order restrictions as an unequivocally non-remnant *precisely*.

The distributional differences between Irish English *all* and Standard American English *exactly/precisely* with respect to the pre-complementizer positioning in (ii) most likely reflects contrasting low-level prosodic conditions on the place-

For the sake of exposition, we start with the analysis of non-floating (93b) and then extend it to the floating *all* in (93a).

(93) a. Who did Frank tell you **all** that they were after __?
b. Who **all** did Frank tell you that they were after __?

For the non-stranded case, we posit the following entry for *all* as a higher-order modifier for a *wh*-operator (mapping a (Q⇂(S⇂NP)) to another (Q⇂(S⇂NP))):

(94) $\lambda\rho\lambda\sigma.\rho(\lambda\varphi.\varphi) \bullet \mathsf{all} \bullet \sigma(\boldsymbol{\epsilon})$; $\lambda\mathscr{F}\lambda P\lambda x_C.\mathscr{F}(P)(x)$; (Q⇂(S⇂NP))⇂(Q⇂(S⇂NP))
defined only if the domain set C for x is above the contextually relevant standard for high precision

This may look somewhat complex, but all it does is impose a certain restriction on the interpretation of the semantic variable x bound by the *wh*-operator. The semantic restriction imposed on x dictates that it be chosen from a domain set (i.e., contextually determined set of individuals) C which counts as sufficiently 'precise' in the context in question. By applying (94) to the *wh*-question operator *who* in (95), we obtain (96), which then licenses the semantics (97) for (93b).

(95) $\lambda\sigma.\mathsf{who} \bullet \sigma(\boldsymbol{\epsilon})$; $\lambda P\lambda x.\mathbf{wh}_{\mathbf{person}}(x)(P)$; Q⇂(S⇂NP)

(96) $\lambda\sigma.\mathsf{who} \bullet \mathsf{all} \bullet \sigma(\boldsymbol{\epsilon})$; $\lambda P\lambda x_C.\mathbf{wh}_{\mathbf{person}}(x)(P)$; Q⇂(S⇂NP)
defined only if the domain set C for x is above the contextually relevant standard for high precision

(97) $\lambda x_C.\mathbf{wh}_{\mathbf{person}}(x)(\mathbf{tell}(\mathbf{you})(\mathbf{after}(x)(\mathbf{they}))(\mathbf{frank}))$
defined only if the domain set C for x is above the contextually relevant standard for high precision

The idea here is that by manipulating the domain set in the direction of increasing precision, things that are normally ignored enter into the domain of entities that the question sentence inquires about. For example, suppose that a police officer is interrogating a witness in an investigation of an issue in which a foregn spy John died after having

---

ment of modifiers of *wh*-words (of different sizes). We therefore assume, following Sag and Levine (2006), that this distributional difference doesn't affect the plausibility of the adverb analysis of Irish English *all*.

lunch with a suspicious person. In this situation, *What **all** did John eat?* is likely a more appropriate (and less ambiguous) question than *What did John eat?*, and it calls for a higher degree of precision and completeness for a proper answer.

Moving on to the floating *all*, we take this *all* to be syntactically a VP adverb which is reordered to the position immediately preceding the complement clause. This can be dealt with by some kind of surface reordering rule governing adverbs (see, e.g., Kubota 2014 for one approach in TLG), and it is motivated by the parallel distribution between *all* and the 'non-*wh*-remnant' adverb *precisely* noted in footnote 14. We can then take the combinatoric structure underlying the matrix VP in (93a) to be something like the following, where x is the free variable corresponding to the embedded gap:

(98) $\mathbf{ALL}_{\mathrm{VP/VP}}(\mathbf{TELL}_{\mathrm{VP/S'/NP}}(\mathbf{YOU}_{\mathrm{NP}})$
$(\mathbf{THAT}_{\mathrm{S'/S}}(\mathbf{WERE}_{\mathrm{VP/VP}}(\mathbf{AFTER}_{\mathrm{VP/NP}}(x_{\mathrm{NP}}))(\mathbf{THEY}_{\mathrm{NP}}))))$

Floating *all* then has the semantics analogous to the non-floating *all* in (94), with the only difference being that in the case of the floating *all*, the semantic variable that it targets is still *unbound* in the term that it takes as its argument as a VP adverb:

(99) $\mathbf{ALL}_{\mathrm{VP/VP}} = \mathsf{all}; \lambda P.P; \mathrm{VP/VP}$
where $\mathbf{ALL}_{\mathrm{VP/VP}}(\alpha)$ is defined only if all elements $x_C \in fv_{\mathrm{X}_{+wh}}(\alpha)$ are such that the domain set C for x is above the contextually relevant standard for high precision

This imposes exactly the same restriction as the non-floating *all* on the variable x that the question operator ranges over. We thus obtain the same final translation for (93a) as for (93b), namely, (97). Thus, though the exact way in which *all* contributes its meaning in the compositional process is somewhat different in the two cases, we effectively get the same result as McCloskey (2000), preserving the key insight of his analysis that there is a tight semantic connection between the *wh*-phrase and the stranded adverb *all*, but doing away with the undermotivated assumption that the latter forms a syntactic unit with the former in the underlying structure and is a movement remnant in the surface structure.

The analysis of the distribution and interpretation of floating *all* in Irish English sketched above takes the 'stranded' *all* to be an

adverb-like operator that targets the denotation of the free variable in the subproof and imposes an additional restriction on its interpretation. Interestingly, at least some of the cases of EPM reported in the literature of the 'remnant movement' type seem to be amenable to a similar treatment. For example, the 'stranding' of quantifier-like elements in Wolof, reported in Torrence 2018 (cited in Davis 2020), consists of a paradigm such as the following:

(100) a. [F-an **f-eeneen**]$_k$ l-a Ayda wax ne l-a-a dem t_k ?
where other COP Ayda say that cop.1sg go
'Where else did Ayda say that I went?'

b. F-an$_k$ l-a-nu foog [t_k **f-eeneen**]$_j$ ne la-a togg-e ceeb t_j ?
where cop.3pl think other that cop.1sg cook rice
'Where else do they think that I cooked rice?'

Here, the 'quantifier-like' element *f-eeneen* that exhibits exceptive interpretation (analogous to English *what else*) restricts the interpretation of the 'trace variable' to things that are not identical to some discourse-salient entity.

A somewhat different pattern is found in Polish, in the following paradigm originally reported by Wiland (2010) (again, we reproduce the data from Davis 2020).

(101) Jaki$_k$ (samochód) Paweł kupił swojej żonie t_k
what car Pawel bought his wife
(samochód)?
car
'What car did Pawel buy his wife?'

(102) a. Jaki$_k$ Paweł kupił [$_{VP}$ [t_k **samochód**]$_j$ swojej żonie t_j]?
what Pawel bought car his wife
'What car did Pawel buy his wife?'

b. Jaki$_k$ Paweł [$_{vP}$ [t_k **samochód**]$_j$ kupił swojej żonie t_j]?
what Pawel car bought his wife
'What car did Pawel buy his wife?'

c. Jaki$_k$ myślisz [$_{CP}$ [t_k **samochód**]$_j$ (*że) Paweł
what think.you car that Pawel
kupił swojej żonie t_j]?
bought his wife
'What car do you think that Pawel bought his wife?'

In these examples, it appears as though the head noun of an extracted *wh*-phrase gets stranded at intermediate landing sites, in an apparent violation of the Left Branch Condition. However, these examples are amenable to a different type of analysis, where the apparently 'stranded' element *samochód* 'car' is again a 'trace-targeting' domain restrictor of some sort, restricting the domain set C to $C \cap$ **car**.

What we can see from the above (including Irish English *all*) is that the fact that some element is semantically related to the *wh*-phrase does not necessarily mean that the expression in question has to form a syntactic unit with the *wh*-phrase at some level of syntactic representation. The alternative analyses we have suggested for these so-called 'remnant stranding' EPM cases crucially exploit the key property of our approach that this phenomenon makes reference to the intermediate status of syntactic derivation/meaning computation involving a hypothetically assumed element. It is interesting to see that items that are 'retooled' for EPM in these languages all have essentially the same semantic function of domain restriction for the targeted variable.

5.3 *A brief note on cyclicity more generally*

Alongside the EPM effects reviewed above, more abstract types of arguments for the notion of cyclicity have been offered in the literature. We review some of these briefly here, with preliminary remarks about their possible implications for our meaning-centered approach. This class of phenomena are potentially important for a comparison between our approach and the standard configurational approaches as they pertain more directly to the architecture of the syntax-semantics interface. In what follows, we discuss in turn (i) arguments involving reflexive binding; (ii) arguments involving the interactions between reconstruction effects in variable binding and Condition C effects and (iii) arguments involving parasitic gap licensing.

First, the binding pattern of the sort exemplified by (103) has sometimes been adduced in the literature in favor of cyclic movement (see, e.g., Barss 2001).

(103) [Which pictures of himself$_{i/j}$] does John$_i$ think t that Bill$_j$ hates t?

The idea here is that the two trace positions in (103) make available reconstruction sites for the fronted *wh*-phrase containing the reflexive, and choosing one or the other satisfies the local c-command requirement with either *John* or *Bill* as the antecedent. But this argument is quite problematic. As den Dikken (2018) notes, the acceptability of (103) on its two readings is compatible with an alternative, non-configurational account for exempt anaphors (of the sort advocated, e.g., by Pollard and Sag (1992) and Reinhart and Reuland (1993)).

A more elaborate type of argument for cyclic movement has been adduced by authors such as Sauerland (1998), involving reconstruction effects in variable binding and Condition C effects. For example, Sauerland notes that the following contrast due to Lebeaux (1992) can be explained by assuming that Condition C applies at LF and that reconstruction of a moved element to an intermediate landing site is possible (for the purpose of variable binding):

(104) a. [Which paper that he$_k$ gave to Mary$_j$]$_i$ did every student$_k$ think t_i' that she$_j$ would like t_i ?

b. *[Which paper that he$_k$ gave to Mary$_j$]$_i$ did she$_j$ think t_i' that every student$_k$ would like t_i ?

In (104b), the fronted *wh*-phrase has to reconstruct to the most deeply embedded trace position t_i for the pronoun it contains to be bound by a c-commanding quantifier. But this incurs a violation of Condition C. By contrast, in (104a), there is an option for the *wh*-phrase to reconstruct to the intermediate trace position t_i', which simultaneously satisfies the variable binding condition for the pronoun and Condition C for the R-expression *Mary*.

While offering a full analysis is beyond the scope of this paper, we sketch a possible approach within a TLG setup. Jäger (2005, 174–178) proposes an analysis of reconstruction effects in TLG that has two key components: (i) binding of pronouns and reflexives is mediated by hypothetical reasoning, and reflected explicitly in the syntactic types of

the binder and the pronoun, adopting Jacobson's (1999) 'pronouns as identity function' approach; (ii) filler-gap linkage transparently preserves the binding relation, by copying the pronoun-containing status of the filler to the 'gap site' via syntactic type encoding of binding (correlating with semantic type).

On this type of approach, the rough form of derivations for the examples in (104) will look like the following:

(105) a. which f (where $f(x)$ is a paper x gave to Mary) [λf. did every student [λy. think (λx[**she** would like x]($\boldsymbol{f}(y)$))(y)]]

b. which f (where $f(x)$ is a paper x gave to Mary) [λf. did **she** think [every student [λy [λz [y would like z]($\boldsymbol{f}(y)$)]]]]

The fronted *wh*-expression receives a functional interpretation involving a person-to-paper mapping f, reflecting its pronoun-containing status. Crucially, this functional variable f (which takes the bound variable y as an argument) has to be introduced in the proof in the most deeply embedded clause in (105b) to enforce binding of the individual variable y by the quantifier. By contrast, in (105a), we can wait till the intermediate clause is built to introduce f since the quantifier appears in the intermediate clause (this is parallel to the availability of the intermediate trace position t_i' in Sauerland's LF-based account). This results in a difference in the structural relationship between the pronoun *she* and the functional variable f (the latter of which gets bound by the fronted *wh*-phrase). Assuming that the R-expression-containing status can be copied from the fronted *wh*-phrase to the f variable (via some feature-matching mechanism, for example) and assuming that Condition C is a condition on the form of the logical proof (parallel to Sauerland's treatment of Condition C as an LF condition), the contrast between (105a) and (105b) follows from the fact that *she* 'c-commands' f in (105b) but not in (105a). While this is still preliminary, it should at least be clear that TLG offers an analysis that preserves the core ideas of Sauerland's LF-based account.

Finally, there is another type of evidence involving parasitic gap licensing due to Nissenbaum (2000).

(106) a. Who did you praise __ to the sky [after criticizing __] [in order to surprise __]?

b. Who did you praise __ to the sky [after criticizing __] [in order to surprise **him**]?

c. *Who did you praise __ to the sky [after criticizing **him**] [in order to surprise __]?

Roughly, the idea is that a moved *wh*-phrase licenses a parasitic gap along the way, in a successive cyclic manner. In (106a), the two adjunct clauses are both inside the largest S hosting the fronted filler, and the *wh*-phrase licenses the gaps inside them as it passes through the stacked vPs. (106b) is different from (106a) in that the outer adjunct clause (*in order to surprise him*) adjoins from outside to a structure in which the filler-gap linkage is completely established. (106c) is the problematic case, in which the offending inner adjunct clause (*after criticizing him*) does not host a gap. The absence of a parasitic gap in the inner adjunct clause prevents cyclic movement of the *wh*-phrase which is required to license the parasitic gap in the outer clause. Nissenbaum takes parasitic gaps to be licensed by an empty operator at LF. This entails that the type of interaction between overt *wh*-movement and parasitic gap licensing in (106) necessitates a 'single cycle' architecture (which abandons the standard T-model) in which overt and covert movement operations are interwoven.

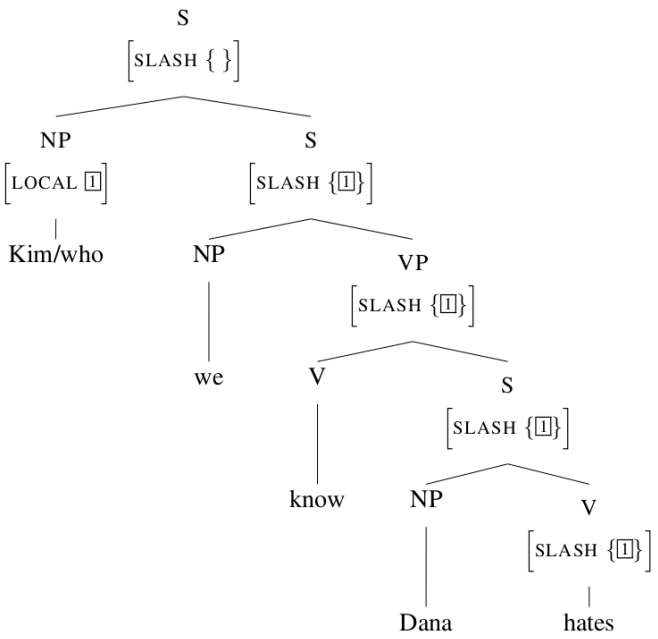
There is an intriguing similarity between the architecture of Hybrid TLG and the 'single cycle' model advocated by Nissenbaum (2000): essentially, Hybrid TLG embodies a 'single cycle' model by design, in that it models overt and covert movement via the same mechanism of prosodic lambda binding within a single model of derivation as logical inference. Interestingly, this architectural design has been independently arrived at without any prior considerations of anything like the Nissenbaum paradigm. This then brings up a question worth exploring in future research: would it be possible to reinterpret the Nissenbaum paradigm within the 'meaning centered' approach we have argued for? Such a reinterpretation would involve viewing both parasitic gap licensing and 'cyclic movement' in processing-oriented terms (the latter along the lines we briefly speculate on at the end of Section 5 on p. 155). We leave this interesting question for future research.

5.4 *Comparison with a feature-percolation analysis of extraction pathway marking in HPSG*

At this point, the key differences between our proof-theoretic analysis and the successive cyclic analysis standard in derivational approaches should be clear. In the syntactic literature, an alternative to the derivational analysis has been proposed by Bouma *et al.* (2001) in the constraint-based framework of HPSG that makes extensive use of the feature percolation mechanism of the framework. We briefly compare our approach with this HPSG approach in this section.

(107) illustrates the HPSG analysis of extraction.

(107)

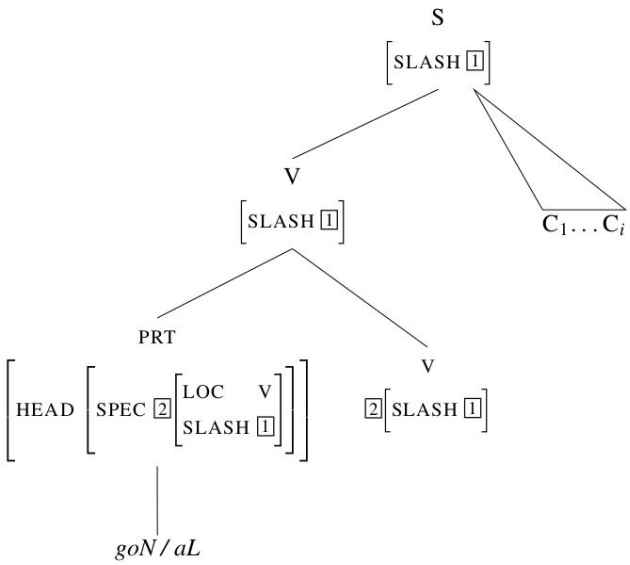


In HPSG, the SLASH feature is employed for indicating that a phrase contains a gap position (in the object of the verb *hates* in (107)). As in (107), this information is successively inherited from daughter to mother via the feature percolation mechanism inherent to HPSG, until the corresponding filler is found (at the top S node).

Given this general analysis of extraction, in the case of complementizer marking in Irish, the choice of the morphological form of the

complementizer can simply be made sensitive to the locally encoded value of the SLASH feature of the verbal projection that *goN/aL* directly combines with, since this feature indicates whether the clause in question contains a gap or not. This is schematically shown in (108).

(108)



The key difference, then, between the TLG analysis and this feature-percolation analysis in HPSG is the following. In the latter, the complementizer choice is dependent on the local syntactic information alone. This is in keeping with the locality condition in HPSG (see, e.g., Sag 2010) and it exploits the general SLASH inheritance mechanism that mediates nonlocal filler-gap linkage via a chain of local feature passing. By contrast, in our TLG analysis, the complementizer choice depends on the existence of an unwithdrawn hypothesis in the subproof (which may be deeply embedded). We have already noted above that this infringes the tenet of direct compositionality in traditional CG, according to which proofs are not representational objects. The reader should now see a connection between HPSG and traditional CG: the CG compositionality thesis roughly corresponds to the locality condition in HPSG – indeed, they are likely to stem from ideas that shaped the common basic form of nonderivational syntactic theories in the 1980s.

While a casual cross-theoretic comparison can be misleading, there does seem to be a tradeoff about which part of the grammar needs to be made complex in the two approaches. Essentially, the HPSG approach abides by the locality principle by slightly enriching the local information encoded at each syntactic node. By contrast, the TLG approach does away with explicit feature percolation at the cost of violating the locality principle in a limited way – limited since all that this approach exploits is a 'filter' constraint that checks the existence of a free variable within a subterm (which conceptually corresponds to the 'tentative assumption' driving hypothetical reasoning in filler-gap linkage).[15] Note that this doesn't involve complex manipulations ('transformations') of the structures of the subterms themselves, or anything that resembles the notion of 'phase' in minimalism (a proof-theoretic analog for this would be a set of meta-constraints imposing an explicit 'control structure' of some sort on proof strategy). In this sense, our proposal is structure-sensitive, but arguably *not* procedural, at least not in the same way that its derivational counterparts (in various avatars of derivational syntax) are.

As a final point of comparison with the constraint-based view of grammar embodied in HPSG, we would like to cautiously bring up possible implications for processing (we ourselves take the competence grammar and the theory of processing to be in principle distinct; see Kubota 2021, Section 5 in this connection). One might initially think that processing-related considerations would favor the local licensing approach embodied in HPSG. However, note that the plausibility of this type of argument largely depends on the assumption that incremental parsing with complex data structures of the sort assumed in HPSG is cognitively realistic. By contrast, TLG embraces a much more indirect relationship between the grammar and processing. That being said, extraction pathway marking formalized as proof structure marking potentially illuminates a possible connection between grammar and processing that has largely been overlooked in the past literature. In proof-theoretic terms, establishing a filler-gap linkage corresponds to withdrawing a hypothesis at a certain point in a proof by finding

[15] In connection to this point, one might recall the discussion from the 'syntax wars' era by proponents of Generative Semantics, e.g., Postal (1972), that global conditions on derivations can always be mimicked by feature marking.

a 'matching' premise (i.e., one that is looking to combine with a conditional statement derived from that hypothesis). Viewing syntactic parsing as proof search – which is a common perspective in TLG – such a complex proof strategy is very likely labor-intensive for the human online parser. It is then not too surprising that some natural languages have developed devices for explicitly flagging the intermediate statuses of the subproofs involved in such proofs, so as to efficiently narrow down the proof search space. Thus, this view offers a particularly natural way of understanding extraction pathway marking as a functionally motivated strategy, one that has fully developed into a grammatically encoded distinction in certain languages.[16]

6 CONCLUSION

We have advocated a new analysis of extraction pathway marking which essentially views this phenomenon as linguistic encoding of proof structure. This has several empirical, technical and conceptual implications that are worth exploring further in future research.

Technically, those familiar with the CG tradition will likely frown on our proposal as it (at least partly) abandons an influential idea of direct compositionality in CG research. We would like to remind such readers that the way our approach makes reference to proof structure is relatively modest, as it merely involves the notion of free variables in a typed lambda calculus (something that is already needed in semantic interpretation anyway). To be sure, global reference to structure is allowed, but we find an analogy to classical Transformational Grammar invoked by one referee somewhat misleading, since, unlike the latter, our approach does not involve arbitrary rewriting of the

[16] One might worry that this discussion on the implications on online processing via proof search might make the relationship between competence and performance obscure and complex in a TLG setting, a point rightly raised by one reviewer. We recognize that this is a legitimate worry, but addressing this important issue fully is a task that we have to leave for future study, in relation to efforts to develop a real processing theory taking some form of TLG as the core component of the competence theory.

structures of already constructed proofs. That being said, we recognize that once this 'Pandora's box' is opened, a question arises as to exactly how much of proof structure reference is allowed and how it is constrained in natural language syntax, an issue we leave for future study. It would also be interesting to see what one can come up with as alternative analyses for EPM within approaches of CG that abide by the notion of direct compositionality more strictly, such as CCG.

Turning to the more conceptual (and empirical) aspects, one might wonder what exactly we gain by this reconceptualization of extraction pathway marking/successive cyclicity. We believe that here the main advantage is that a new, meaning-centered approach to the typology of extraction pathway marking comes into sight, which can be contrasted with the more traditional structure-driven approach that has been dominant in the literature. An almost immediate consequence of our approach is that extraction pathway marking makes reference to the *semantic* relationship between an unwithdrawn hypothesis (corresponding to a free variable) and a larger expression containing it. And there are a couple of 'obvious' choices for encoding such semantic sensitivity in specific morpho-syntactic devices, all attested in one language or another:

- **Direct morpho-syntactic EPM marking** (Irish complementizer selection, Belfast English inversion): This is the most straightforward strategy, in which the language marks the extraction pathway on some functional expression that takes a proposition-denoting constituent as an argument, and signals that the latter involves an incomplete proof.
- **EPM via domain restriction on 'trace' interpretation** (Irish English *all* stranding, Dinka plural marking, Wolof Q-like particle, Polish stranded head N): Impose a restriction pertaining to the semantic interpretation of the relevant free variable. Interestingly, this option seems to allow for more word order freedom than the above morpho-syntactic strategy. This may be due to the fact that domain restrictors are not proposition-taking functions but expressions that are originally part of the (extracted) NP or adverbial elements diachronically.
- **EPM via 'information packaging'** (Dinka V2 word order): This is the most abstract and subtle type of encoding in which the 'distin-

guished' status of the free variable (to be bound by some operator in a higher clause) competes for discourse-oriented prominence. Here again, the semantic interpretation of the variable within the subexpression in which it occurs plays a crucial role in licensing the relevant intermediate proof.

These patterns are of course all well-known, but so far as we are aware, the previous literature does not offer a clear answer to the question of *why* EPM often exhibits sensitivity to the interpretation of the semantic variable with respect to the syntactic context in which it appears. Of course a lot more work needs to be done to investigate this typological literature, but we think that our approach is interesting as it has the potential of shedding a new light on this cross-linguistic typology.

To put the present proposal in a still larger context, it is useful to reflect on the larger goals of comparative syntax in the generative tradition. A core idea behind generative comparative syntax is that the combinatoric system underlying syntax has unique properties characterizing human language. Successive cyclicity has been one major (and quite attractive) candidate for such a property. But a logical reconceptualization of this notion we have attempted in this paper leads to a somewhat different perspective: in our TLG analysis, extraction pathway marking reduces to nothing more than a surface manifestation of an intermediate status of a proof. Our conclusion (and contention), then, is simple: cyclicity may initially look like the best candidate for an unreducible *unique* property of human language, but upon closer inspection, it turns out to be a reflection of a *general* property of logic underlying that system.

ACKNOWLEDGMENTS

We'd like to thank Michael Moortgat, Joanna Nykiel, Shinya Okano and Mizuki Sakamoto for useful discussion. The paper has also benefitted from substantial feedback from the reviewers for *JLM*, and we'd like to thank all of those who were involved in the reviewering process at one stage or another. This work was supported by JSPS KAKENHI Grant Number 21K00541 and the NINJAL collaborative research project 'Toward a Computationally-Informed Theoretical Linguistics', and was carried out as part of a project at the Advanced Language Science (E3P) Research Center at NINJAL.

REFERENCES

Kazimierz AJDUKIEWICZ (1935), Die syntaktische Konnexität, *Studia Philosophica*, 1:1–27, English translation in Storrs MCCALL, editor (1967), *Polish Logic: 1920–1939*, pp. 207–231, Oxford University Press, Oxford.

Doug ARNOLD and Danièle GODARD (2021), Relative clauses in HPSG, in Müller *et al.* (2021), pp. 595–663, `https://langsci-press.org/catalog/book/259`.

Yehoshua BAR-HILLEL (1953), A quasi-arithmetic notation for syntactic descriptions, *Language*, 29:47–58.

Matthew BARROS and Robert FRANK (2023), Attention and locality: On clause-boundedness and its exceptions in multiple sluicing, *Linguistic Inquiry*, 54(4):649–684.

Andrew BARSS (2001), Syntactic reconstruction effects, in Mark BALTIN and Chris COLLINS, editors, *The handbook of contemporary syntactic theory*, pp. 670–696, Blackwell, Oxford.

Cedric BOECKX (2012), *Syntactic islands*, Cambridge University Press, Cambridge.

Olivier BONAMI, Danièle GODARD, and Jean-Marie MARANDIN (1999), Constituency and word order in French subject inversion, in Gosse BOUMA, Erhard HINRICHS, Geert-Jan M. KRUIFT, and Richard OEHRLE, editors, *Constraints and resources in natural language syntax and semantics*, pp. 21–40, CSLI, Stanford.

Robert BORSLEY (2010), An HPSG approach to Welsh unbounded dependencies, in Stefan MÜLLER, editor, *Proceedings of the 17th International Conference on Head-Driven Phrase Structure Grammar*, pp. 80–100, CSLI, Stanford.

Gosse BOUMA, Rob MALOUF, and Ivan A. SAG (2001), Satisfying constraints on extraction and adjunction, *Natural Language and Linguistic Theory*, 19(1):1–65.

Joan BRESNAN and Sam A. MCHOMBO (1995), The lexical integrity principle: evidence from Bantu, *Natural Language and Linguistic Theory*, 13(2):181–254.

Wojciech BUSZKOWSKI (1987), The logic of types, in Jan SRZEDNICKI, editor, *Initiatives in logic*, pp. 180–206, Springer, Dordrecht.

Bob CARPENTER (1997), *Type-logical semantics*, MIT Press, Cambridge, MA.

Rui P. CHAVES and Michael T. PUTNAM (2020), *Unbounded dependency constructions: Theoretical and experimental perspectives*, Oxford University Press, Oxford.

Noam CHOMSKY (1973), Conditions on transformations, in Stephen A. ANDERSON and Paul KIPARSKY, editors, *A Festschrift for Morris Halle*, pp. 232–286, Holt, Rinehart and Winston, New York.

Noam CHOMSKY (1977), On *wh*-movement, in Peter CULICOVER, Thomas WASOW, and Adrian AKMAJIAN, editors, *Formal syntax*, pp. 71–132, Academic Press, New York.

Noam CHOMSKY (1981), *Lectures on government and binding*, Foris, Dordrecht.

Noam CHOMSKY (1986), *Barriers*, MIT Press, Cambridge, MA.

Noam CHOMSKY (2000), Minimalist inquiries: The framework, in Roger MARTIN, David MICHAELS, and Juan URIAGEREKA, editors, *Step by step: Essays on Minimalist Syntax in honor of Howard Lasnik*, pp. 89–156, MIT Press, Cambridge, MA.

Sandra CHUNG (1982), Unbounded dependencies in Chamorro grammar, *Linguistic Inquiry*, 13:39–77.

Barbara CITKO (2014), *Phase theory*, Cambridge University Press, Cambridge.

George N. CLEMENTS, James MCCLOSEY, Joan MALING, and Annie ZAENEN (1983), String-vacuous rule application, *Linguistic Inquiry*, 14:1–17.

Robin COOPER (1983), *Quantification and syntactic theory*, volume 21 of *Synthese Language Library*, Reidel, Dordrecht.

Colin DAVIS (2020), Crossing and stranding at edges: On intermediate stranding and phase theory, *Glossa*, 5(1):17.

Philippe DE GROOTE (2001), Towards abstract categorial grammars, in *Association for Computational Linguistics, 39th Annual Meeting and 10th Conference of the European Chapter*, pp. 148–155, Association for Computational Linguistics.

Paul D. DEANE (1992), *Grammar in mind and brain*, de Gruyter, Berlin.

Hans DEN BESTEN (2010), Is there preposition stranding in COMP in Afrikaans? No way!, in Jan-Wouter ZWART and Mark DE VRIES, editors, *Structure preserved: Studies in syntax for Jan Koster*, pp. 57–64, Amsterdam: John Benjamins.

Marcel DEN DIKKEN (2009), Arguments for successive-cyclic movement through SpecCP: A critical review, in *Linguistic Variation Yearbook*, pp. 89–126, John Benjamins.

Marcel DEN DIKKEN (2017), Overtly marked *wh*-paths, in Martin EVERAERT and Henk C. VAN RIEMSDIJK, editors, *The Wiley Blackwell companion to syntax, Second Edition*, John Wiley & Sons.

Marcel DEN DIKKEN (2018), *Dependency and directionality*, Cambridge University Press, Cambridge, UK.

Mark DONOHUE and Anna MACLACHLAN (1999), What agreement in Chamorro?, in Carolyn SMALLWOOD and Catherine KITTO, editors, *Proceedings of AFLA VI*, volume 16, pp. 121–132, University of Toronto Department of Linguistics, Canada: Toronto.

David DOWTY (2007), Compositionality as an empirical problem, in Chris BARKER and Pauline JACOBSON, editors, *Direct compositionality*, pp. 23–101, Oxford University Press, Oxford.

David R. DOWTY, Robert WALL, and Stanley PETERS (1981), *Introduction to Montague semantics*, Reidel Publishing Co., Dordrecht.

Hans DU PLESSIS (1977), *Wh* movement in Afrikaans, *Linguistic Inquiry*, 8:723–726.

Samuel EPSTEIN and T. Daniel SEELY (2002), Rule applications as cycles in a level-free syntax, in Samuel EPSTEIN and T. Daniel SEELY, editors, *Derivation and explanation in the Minimalist Program*, pp. 65–89, Blackwell, Oxford.

Gerald GAZDAR, Geoffrey PULLUM, and Ivan SAG (1982), Auxiliaries and related phenomena in a restrictive theory of grammar, *Language*, 58:591–638.

Thomas GRANO and Howard LASNIK (2018), How to neutralize a finite clause boundary: Phase theory and the grammar of bound pronouns, *Linguistic Inquiry*, 49(3):465–499.

Irene HEIM and Angelika KRATZER (1998), *Semantics in Generative Grammar*, Blackwell, Oxford.

Alison HENRY (1995), *Belfast English and Standard English*, Oxford University Press, Oxford.

Wolfram HINZEN (2012), The philosophical significance of Universal Grammar, *Language Sciences*, 34(5):635–649.

Philip HOFMEISTER and Ivan A. SAG (2010), Cognitive constraints and island effects, *Language*, 86(2):366–415.

William A. HOWARD (1969), The formulae-as-types notion of construction, in Jonathan P. Seldin and J. Roger Hindley (eds.) 1980, *To H. B. Curry: Essays on combinatory logic, lambda calculus, and formalism*, 479–490 (New York: Academic Press).

C.-T. James HUANG (1982), *Logical relations in Chinese and the theory of grammar*, Ph.D. thesis, MIT.

Thomas E. HUKARI and Robert D. LEVINE (1995), Adjunct extraction, *Journal of Linguistics*, 31:195–226.

Pavel IOSAD (2023), Mutation in Celtic, in Peter ACKEMA, Sabrina BENDJABALLAH, Eulàlia BONET, and Antonio FÁBREGAS, editors, *The Wiley Blackwell companion to morphology*, pp. 1565–1606, Wiley-Blackwell, United Kingdom.

Pauline JACOBSON (1999), Towards a variable-free semantics, *Linguistics and Philosophy*, 22(2):117–184.

Gerhard JÄGER (2005), *Anaphora and Type-Logical Grammar*, Springer, Berlin.

Richard KAYNE and Jean-Yves POLLOCK (1978), Stylistic inversion, successive cyclicity and Move NP in French, *Linguistic Inquiry*, 9:595–621.

Robert KLUENDER (1992), Deriving island constraints from principles of predication, in Helen GOODLUCK and Michael ROCHEMONT, editors, *Island constraints: Theory, acquisition, and processing*, pp. 223–258, Kluwer, Dordrecht.

Robert KLUENDER (1998), On the distinction between strong and weak islands: A processing perspective, in Peter CULICOVER and Louise MCNALLY, editors, *The limits of syntax*, volume 29 of *Syntax and Semantics*, pp. 241–279, Academic Press, San Diego.

Yusuke KUBOTA (2014), The logic of complex predicates: A deductive synthesis of 'argument sharing' and 'verb raising', *Natural Language and Linguistic Theory*, 32(4):1145–1204.

Yusuke KUBOTA (2021), HPSG and categorial grammar, in Müller *et al.* (2021), pp. 1331–1394, `https://langsci-press.org/catalog/book/259`.

Yusuke KUBOTA and Robert LEVINE (2020), *Type-logical syntax*, MIT Press, Cambridge, MA.

Yusuke KUBOTA and Robert LEVINE (to appear), The logic of the English auxiliary system, in *Logical Aspects of Computational Linguistics 2021*, Springer, Heidelberg, `https://ling.auf.net/lingbuzz/006040`.

David LEBEAUX (1992), Relative clauses, licensing, and the nature of the derivation, in Susan ROTHSTEIN and Margaret SPEAS, editors, *Perspectives on phrase structure: Heads and licensing*, pp. 209–239, Academic Press, New York.

Julie Anne LEGATE (1998), Verb phrase types and the notion of a phase, unpublished manuscript, MIT.

Robert D. LEVINE and Thomas E. HUKARI (2006), *The unity of unbounded dependency constructions*, CSLI, Stanford.

Christopher D. MANNING, Ivan A. SAG, and Masayo IIDA (1999), The lexical integrity of Japanese causatives, in Robert LEVINE and Georgia GREEN, editors, *Studies in contemporary phrase structure grammar*, pp. 39–79, Cambridge University Press, Cambridge.

Ora MATUSHANSKY (2005), Going through a phase, in Martha MCGINNIS and Norvin RICHARDS, editors, *MIT Working Papers in Linguistics 49: Perspectives on Phases*, pp. 157–181, MIT Press, Cambridge, MA.

James MCCLOSKEY (1979), *Transformational syntax and model-theoretic semantics*, Dordrecht: Reidel.

James MCCLOSKEY (1990), Resumptive pronouns, Ā binding and levels of representation in Irish, in Randall HENDRICK, editor, *The Syntax of the modern Celtic languages*, pp. 199–248, New York: Academic Press.

James MCCLOSKEY (2000), Quantifier float and *wh*-movement in an Irish English, *Linguistic Inquiry*, 31:57–84.

James MCCLOSKEY (2002), Resumption, successive cyclicity, and the locality of operations, in Samuel EPSTEIN and T. Daniel SEELY, editors, *Derivation and explanation in the Minimalist Program*, pp. 184–226, Wiley.

Vedrana MIHALIČEK and Carl POLLARD (2012), Distinguishing phenogrammar from tectogrammar simplifies the analysis of interrogatives, in Philippe DE GROOTE and Mark-Jan NEDERHOF, editors, *Formal Grammar 2010/2011*, pp. 130–145, Springer, Heidelberg.

Michael MOORTGAT (2011), Categorial type logics, in Johan VAN BENTHEM and Alice TER MEULEN, editors, *Handbook of logic and language*, pp. 95–179, Elsevier, Amsterdam, 2 edition.

Michael MOORTGAT (2014), Typelogical grammar, in Edward N. ZALTA, editor, *The Stanford Encyclopedia of Philosophy (Spring 2014 Ed.)*, Stanford University, available at `https://plato.stanford.edu/archives/spr2014/entries/typelogical-grammar/`.

Glyn MORRILL (1994), *Type Logical Grammar: Categorial logic of signs*, Kluwer, Dordrecht.

Gereon MÜLLER (2010), On deriving CED effects from the PIC, *Linguistic Inquiry*, 41(1):35–82.

Stefan MÜLLER, Anne ABEILLÉ, Robert D. BORSLEY, and Jean-Pierre KOENIG, editors (2021), *Head-Driven Phrase Structure Grammar: The handbook*, Language Science Press, Berlin, `https://langsci-press.org/catalog/book/259`.

Reinhard Muskens (2003), Language, lambdas, and logic, in Geert-Jan Kruijff and Richard Oehrle, editors, *Resource sensitivity in binding and anaphora*, pp. 23–54, Kluwer, Dordrecht.

Frederick J. Newmeyer (2016), Nonsyntactic explanations of island constraints, *Annual Review of Linguistics*, 2:187–210.

Jon Nissenbaum (2000), *Investigations of covert phrase movement*, Ph.D. thesis, MIT.

Richard T. Oehrle (1994), Term-labeled categorial type systems, *Linguistics and Philosophy*, 17(6):633–678.

Dennis Ott (2014), Review of Cedric Boeckx, *Syntactic Islands* (CUP, 2012), *Language*, 90(1):287–291.

Carl Pollard and Ivan A. Sag (1992), Anaphors in English and the scope of binding theory, *Linguistic Inquiry*, 23(2):261–303.

Paul Postal (1972), A global constraint on pronominalization, *Linguistic Inquiry*, 3(1):35–59.

Geoffrey K. Pullum (1992), The origins of the cyclic principle, in J. M. Denton, G. P. Chan, and C. Canakis, editors, *CLS28: Papers from the 28th Regional Meeting of the Chicago Linguistic Society*, volume 2, pp. 209–236, Chicago Linguistic Society.

Tanya Reinhart and Eric Reuland (1993), Reflexivity, *Linguistic Inquiry*, 24(4):657–720.

John Robert Ross (1967), *Constraints on variables in syntax*, Ph.D. thesis, MIT, reproduced by the Indiana University Linguistics Club.

Ivan Sag, Rui P. Chaves, Anne Abeillé, Bruno Estigarribia, Frank Van Eynde, Dan Flickinger, Paul Kay, Laura Michaelis, Stefan Müller, Geoffrey Pullum, and Thomas Wasow (2020), Lessons from the English auxiliary system, *Journal of Linguistics*, 56:87–155.

Ivan A. Sag (2010), Feature geometry and predictions of locality, in Anna Kibort and Greville G. Corbett, editors, *Features: Perspectives on a key notion in linguistics*, pp. 236–271, Oxford University Press, Oxford.

Ivan A. Sag and Robert D. Levine (2006), Irish English and the status of intermediate traces, paper presented at the 2006 meetings of the LSA; available at https://www.asc.ohio-state.edu/levine.1/UE.pdf.

Uli Sauerland (1998), *The meaning of chains*, Ph.D. thesis, MIT.

Peter Sells (1984), *Syntax and semantics of resumptive pronouns*, Ph.D. thesis, University of Massachusetts Amherst.

Maziar Toosarvandani (2016), Embedding the antecedent in gapping: Low coordination and the role of parallelism, *Linguistic Inquiry*, 47(2):381–390.

Harold TORRENCE (2018), Left edge agreeing elements in Wolof wh-questions, unpublished manuscript, University of California Los Angeles.

Coppe VAN URK (2020), Successive cyclicity and the syntax of long-distance dependencies, *Annual Review of Linguistics*, 6:111–130.

Coppe VAN URK and Norvin RICHARDS (2015), Two components of long-distance extraction: Successive cyclicity in Dinka, *Linguistic Inquiry*, 46:113–155.

Heinrich WANSING (1992), Formulas-as-types for a hierarchy of sublogics of intuitionistic propositional logic, in David PEARCE and Heinrich WANSING, editors, *Nonclassical logics and information processing*, pp. 125–145, Springer, Berlin and Heidelberg.

Bartosz WILAND (2010), Overt evidence from left-branch extraction in Polish for punctuated paths, *Linguistic Inquiry*, 41:335–347.

Annie ZAENEN (1983), On syntactic binding, *Linguistic Inquiry*, 14(3):469–504.

Yusuke Kubota
0000-0002-8468-5857
kubota@ninjal.ac.jp

NINJAL

Robert Levine
0000-0001-8938-1839
levine.1@osu.edu

Ohio State University

Yusuke Kubota and Robert Levine (2025), Against successive cyclicity: A proof-theoretic account of extraction pathway marking, *Journal of Language Modelling*, 13(1):99–165
https://dx.doi.org/10.15398/jlm.v13i1.339