Type-Logical Syntax

OA Funding Provided By:

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

The title-level DOI for this work is:

doi:10.7551/mitpress/11866.001.0001

7 Filler-Gap Dependency

In this chapter, we extend the analysis of long-distance dependencies sketched briefly in chapter 2. The main purpose of this extension is twofold. First, there are certain empirical phenomena discussed extensively in the literature of phrase structure–based syntactic theories under the rubric of "syntactic binding domain phenomena" or "extraction pathway marking" (e.g., Zaenen 1983; Levine and Hukari 2006, 119; Chaves 2009, 48), which initially seem to pose significant difficulties to the type of analysis of extraction in terms of hypothetical reasoning from chapter 2, whose key analytic idea goes back to Muskens (2003). Though a complete analysis of this rather intricate empirical domain (which involves several languages with well-attested patterns of this sort) is beyond the scope of the present monograph, we find it important to skech the core idea of a possible elaboration of the Muskens-style analysis of extraction in TLCG. The account we formulate here should not be understood as a full-fledged analysis but rather as a proof of concept that the logical deductive system underlying our approach can be augmented with devices (which have empirical motivation in other domains) to deal with the locality effects exhibited by the extraction pathway marking data.

The second reason for extending the fragment of long-distance dependencies here is that the central focus of the next chapter is an interaction between ellipsis and extraction, and for this reason, we need to broaden the coverage of the simple analysis of extraction in chapter 2 so that it can deal with a wider range of data. In particular, an explicit analysis of parasitic gap licensing is needed in order to deal with the parasitic gap licensing by elided VPs discussed by Postal (1993), Shimada (1999), and Kennedy (2003). Since multiple-gap phenomena (including parasitic gaps) pose a potentially deep foundational issue for TLCG, we do not mean to offer here a definitive answer to the question of how best to deal with multiple gaps in TLCG. Our goal here is more modest. We sketch an outline of one possible approach that does not involve a radical reworking of the foundations of the architecture of the syntax-semantics interface. We then show, in the next chapter, that this analysis interacts in a simple and direct fashion with the analysis of VP ellipsis from chapter 6 to straightforwardly predict the ellipsis/extraction interaction data that Kennedy (2003) adduces for his deletion-based analysis of VP ellipsis in English.

It may ultimately turn out that the specific analyses of extraction pathway marking and multiple-gap phenomena we introduce in this chapter need to be replaced by more elaborate/sophisticated analyses, possibly together with some major reformulations of (at least part of) the logical deductive system itself, for example, along the lines of a recent proposal by Morrill (2017) for the analysis of multiple gaps. If this turns out to be the case, the specific analysis we formulate below should be viewed as a touchstone that any analysis that supersedes it should be able to account for, with comparable or (ideally) less extra machinery involved. In any event, the treatment of long-distance dependencies is a relatively underdeveloped domain of inquiry in TLCG research, and for this reason, our own investigations in this chapter have a more exploratory nature than the material presented in previous chapters. Where there are open issues and alternative analytic perspectives, we note them as such in the ensuing discussions.

7.1 Extraction Pathway Marking

7.1.1 The Reality of Extraction Pathway Marking

There is a basis for viewing extraction as a series of locally constructed linkages: the existence of languages which distinctively mark in some way the syntactic domains intervening between fillers and their associated gaps. This pattern, first identified as a cross-linguistic pattern in Zaenen (1983), immediately confronts accounts such as Muskens's (2003) with a descriptive difficulty which, unlike island effects, is unlikely to originate in non-syntactic factors. The problem is that, on Muskens's account, the morpho-syntactic/prosodic "registration" of a filler-gap pathway will not be possible. In Muskens's analysis of extraction, the linkage between the gap and filler is established via a single step of hypothetical reasoning via the vertical slash, and given the nature of hypothetical reasoning, whether or not a particular linguistic expression "contains a gap" is impossible to tell since a hypothesis does not necessarily correspond to a gap within the overall architecture of grammar.

Consider, for concreteness, the facts first reported by James McCloskey (1979) for Irish:

(335) Shíl mé **goN** mbeadh sé ann.
thought I COMP would-be he there
'I thought that he would be there.'

(336) Dúirt mé **gurL** shíl me **goN** mbeadh sé ann.
said I **goN**+PAST thought I **goN** would-be he there
'I said that I thought that he would be there.'

In (335)–(336), we see multiple instances of *g*- complementizers, as is appropriate to a series of clauses in which all valence requirements are locally satisfied by overt constituents. In (337), however, while the lower clause contains no gap "sites" and is appropriately marked with a *g*- complementizer, the upper clause, which is missing its subject (as indicated by the gap marker __), is identified via an *a*- form.[1]

(337) an fear **aL** shíl __ **goN** mbeadh sé ann
the man COMP thought __ COMP would-be he there
'the man that thought he would be there'

Finally, the examples in (338)–(339) display the characteristic local registration in Irish of the linkage between the filler and the gap over an arbitrary number of structural levels.

(338) an fear **aL** shíl mé **aL** bheadh __ ann
the man COMP thought I COMP would-be __ there
'the man that I thought would be there'

(339) an fear **aL** dúirt mé **aL** shíl mé **aL** bheadh __ ann
the man COMP said I COMP thought I COMP would-be __ there
'The man that I said I thought would be there'

Moreover, 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 *g*- class complementizers, a point illustrated in (337) and at still greater structural depth in (340).

(340) an fear **aL** dúirt sé **aL** shíl __ **goN** mbeadh sé ann
the man COMP said he COMP thought __ COMP would-be he there
'the man that he said thought he would be there'

Apart from its inherent interest as a convincing demonstration of the local nature of information sharing in extraction dependencies, the grammatical marking of extraction pathways in Irish provides further corroboration of the convergence between valent and adjunct extraction.

Note first that Irish supports the argument that extraction of adjuncts (corresponding to functors in our framework) is mediated by exactly the same mechanisms as are required for the extraction of arguments:

(341) an lá **aL** bhí muid nDoire __
[the day]$_j$ COMP were we in Derry e$_j$

1. The locution "gap site" is used here as an informal abbreviation for "string position in which a phonologically overt substring would normally appear in the saturation of some functor type."

'the day we were in Derry'

(342) Cén uair **aL** tháinig siad na bhaile __
[which time]$_j$ COMP came they home e$_j$
'What time did they come home?'

Neither the copula in Irish nor *tháinig* selects a constituent corresponding to a temporal description, but in both of these cases we find a fronting of these modificatory phrases associated with one or more *a-* series complementizers. Similarly, *béarfaí* in (343) does not select a constituent indicating location, but when an adjunct is extracted from a multiply-embedded position, all the clause boundaries between the extraction site and the filler are marked by the *a-* complementizer.

(343) I mBetlehem **aL** dúirt na targaireachtaí **aL** béarfaí an
in Bethlehem COMP said the prophecies COMP would-be-born the
Slánaitheoir __
Savior __
'It was in Bethlehem that the prophecies said the Savior would be born.'

The data in (341)–(343) establish conclusively that the displacement of adjuncts is mediated by the same mechanism as the displacement of material selected by a head.

But this family of morpho-syntactic patterns presents a major challenge to the Muskens treatment of filler-gap dependencies. Since the history of proof does not correspond to a representational object expressing the description of natural language sentences, there is no obvious locus for the identification of the notion "connectivity pathway" that would be a direct analogue for the appearance of traces in a chain of functional projections at clause boundaries in P&P formulations of these pathways or the appearance of nonempty SLASH specifications in various avatars of phrase structure grammar, starting with Gazdar et al. (1985) and maintained and applied to the marking of extraction locality in, e.g., Bouma et al. (2001). Since the order of proof steps corresponding to the embedding of material via functor saturations has no comparable standing in the representation language of TLCG, we seem to be at a loss for resources applicable to capturing patterns such as the *goN/aL* alternations in (335)–(342).

7.1.2 Marking Extracted Arguments with a Syntactic Feature

We first elaborate the analysis of extraction from chapter 2 by adding a syntactic feature for distinguishing extracted and overt phrases. The extraction operator is thus revised to specifically require the missing NP to carry the +*wh* feature:

(344) $\lambda\sigma.\mathsf{what} \circ \sigma(\boldsymbol{\varepsilon}); \mathbf{wh}(\mathbf{obj}); \mathrm{Q}{\upharpoonright}(\mathrm{S}{\upharpoonright}\mathrm{NP}_{+wh})$

This modest elaboration has an apparently significant technical consequence (which, however, turns out to be a nonissue). Consider a very simple sentence such as (345):

(345) I wonder what John ate.

In order to supply a variable whose abstraction can feed the extraction operator (344), we need to saturate *ate* with an NP_{+wh} variable. But *eat* can combine with other NPs in non-extraction constructions as well, making the default statement of this verb's syntactic type $\mathrm{VP}/(\mathrm{NP}_{+wh} \vee \mathrm{NP}_{-wh})$.[2]

Disjunctive categories have their own inference rules—duals, in a specific sense, of the rules for the conjunctive type constructor $\wedge$ discussed in section 6.3—which are independently motivated empirically in the analysis of feature neutralization effects, as discussed in detail in Bayer (1996). We reproduce here the $\vee$ Introduction rules from Bayer (1996):[3]

(346) a. Left Join Introduction

$$\frac{a; \mathcal{F}; B}{a; \mathcal{F}; A \vee B}\vee\mathrm{I}_l$$

b. Right Join Introduction

$$\frac{a; \mathcal{F}; A}{a; \mathcal{F}; A \vee B}\vee\mathrm{I}_r$$

Using these rules, we can directly prove the lemma (or theorem) $\mathrm{VP}/(\mathrm{NP}_{+wh} \vee \mathrm{NP}_{-wh}) \vdash \mathrm{VP}/\mathrm{NP}_{+wh}$. The proof is straightforward, completely cognate to that for the elementary theorem in classical propositional logic $(\phi \vee \psi) \supset \rho \vdash \phi \supset \rho$:

2. We use the notation $\mathrm{NP}_{\#wh}$ as an abbreviation for $\mathrm{NP}_{+wh} \vee \mathrm{NP}_{-wh}$ throughout the rest of this chapter and chapter 8. Since NP itself is an abbreviation for a fully specified syntactic category in which all feature values are fully specified (cf. A.4 in the appendix), we need to be more precise about what we actually mean by saying that $\mathrm{NP}_{\#wh}$ abbreviates $\mathrm{NP}_{+wh} \vee \mathrm{NP}_{-wh}$. What we mean here is that the symbol "NP" in the abbreviatory notation $\mathrm{NP}_{\#wh}$ for the disjunctively specified category is to be understood as a *fully specified* NP category (modulo the *wh* feature), such as $\mathrm{NP}_{+p,sg}$ (assuming that, aside from *wh*, *p* and *sg*/*pl* are the only syntactic features appropriate for NP). Given this assumption, the abbreviatory notation for the transitive verb $(\mathrm{NP}_{\#wh}\backslash\mathrm{S})/\mathrm{NP}_{\#wh}$ can, for example, be instantiated either as (ia) or (ib), but *not* as (ic) or (id):

(i) a. $((\mathrm{NP}_{+p,sg,+wh} \vee \mathrm{NP}_{+p,sg,-wh})\backslash\mathrm{S})/(\mathrm{NP}_{+p,pl,+wh} \vee \mathrm{NP}_{+p,pl,-wh})$
b. $((\mathrm{NP}_{-p,sg,+wh} \vee \mathrm{NP}_{-p,sg,-wh})\backslash\mathrm{S})/(\mathrm{NP}_{+p,pl,+wh} \vee \mathrm{NP}_{+p,pl,-wh})$
c. $((\mathrm{NP}_{+p,pl,+wh} \vee \mathrm{NP}_{+p,sg,-wh})\backslash\mathrm{S})/(\mathrm{NP}_{+p,pl,+wh} \vee \mathrm{NP}_{+p,pl,-wh})$
d. $((\mathrm{NP}_{+p,sg,+wh} \vee \mathrm{NP}_{+p,sg,-wh})\backslash\mathrm{S})/(\mathrm{NP}_{+p,pl,+wh} \vee \mathrm{NP}_{-p,pl,-wh})$

In (ic), the subject NP has conflicting number specifications, and in (id), the value for the *p* (pronominal) feature for the object NP doesn't match.

3. Bayer assumes that terms inhabiting conjoined and disjoined types cannot differ in their semantics, regardless of which subtype they belong to. Such signs, in his terminology, reflect a semantically "nonpotent" interpretation of the meet and join connectives, in contrast to Morrill's (1994) treatment, in which $X \vee Y$ can combine different semantics for X, Y to yield ordered-pair interpretations. We follow Bayer's analysis for join, taking the semantics of disjoined types to be nonpotent—but, per our analysis of examples such as (316) (from chapter 6), we assume the meet connective to be associated with a semantically potent denotation, whose component meanings are obtained along the lines outlined in Carpenter (1997) for the treatment of "sum" types.

(347)
$$\cfrac{\cfrac{k;\ \mathrm{VP}/(\mathrm{NP}_{+wh}\vee\mathrm{NP}_{-wh}) \qquad \cfrac{[\varphi_0;\mathrm{NP}_{+wh}]^1}{\varphi_0;\ \mathrm{NP}_{+wh}\vee\mathrm{NP}_{-wh}}\,{\vee}\mathrm{I}}{k\circ\varphi_0;\ \mathrm{VP}}\,/\mathrm{E}}{k;\ \mathrm{VP}/\mathrm{NP}_{+wh}}\,/\mathrm{I}^1$$

This proof can be trivially generalized schematically in the form of $X\|Y \vdash X\|Z$ (where $Z \vdash Y$), with X, Y, Z variables over syntactic types and $\|$ a variable over implicational connectives (i.e., $/$, $\backslash$, and $\upharpoonright$). In writing proofs, we suppress the steps explicitly shown in (347) and pretend as if NPs with the *wh* feature specified (either as $+$ or $-$) can directly saturate $\mathrm{NP}_{\#wh}$ arguments of verbs, as in the following derivation:

(348)
$$\cfrac{\cfrac{\cfrac{\cfrac{\mathsf{ate};\ \mathbf{eat};\ \mathrm{VP}/\mathrm{NP}_{\#wh} \quad [\varphi_0;x;\mathrm{NP}_{+wh}]^1}{\mathsf{ate}\circ\varphi_0;\ \mathbf{eat}(x);\ \mathrm{VP}} \quad \begin{array}{l}\mathsf{john};\\ \mathbf{j};\mathrm{NP}_{-wh}\end{array}}{\mathsf{john}\circ\mathsf{ate}\circ\varphi_0;\ \mathbf{eat}(x)(\mathbf{j});\ \mathrm{S}}}{\lambda\varphi_0.\mathsf{john}\circ\mathsf{ate}\circ\varphi_0;\ \lambda x.\mathbf{eat}(x)(\mathbf{j});\ \mathrm{S}\upharpoonright\mathrm{NP}_{+wh}}\,{\upharpoonright}\mathrm{I}^1 \quad \begin{array}{l}\lambda\sigma.\mathsf{what}\circ\sigma(\boldsymbol{\varepsilon});\\ \mathbf{wh}(\mathbf{obj});\\ \mathrm{Q}\upharpoonright(\mathrm{S}\upharpoonright\mathrm{NP}_{+wh})\end{array}}{\mathsf{what}\circ\mathsf{john}\circ\mathsf{ate}\circ\boldsymbol{\varepsilon};\ \mathbf{wh}(\mathbf{obj})(\lambda x.\mathbf{eat}(x)(\mathbf{j}));\ \mathrm{Q}}$$

7.1.3 Modeling Extraction Pathway Marking

We now assume a systematic encoding of "level of clausal embedding" by means of a numerical index attached to syntactic categories to keep track of the "extraction pathway" of expressions extracted out of embedded positions. For this purpose, we follow the approach of Pogodalla and Pompigne (2012), who employ this type of feature encoding (formally implemented in terms of subtypes in their formulation) for the purpose of syntactically encoding scope islands.[4] We thus assume the following lexicon for Irish:

(349) a. an; $\lambda P.\iota(P)$; $\mathrm{NP}^n_{-wh}/\mathrm{N}^n$
b. dól; $\lambda x\lambda y.\mathbf{drink}(y)(x)$; $(\mathrm{S}^n/\mathrm{NP}^n_{\#wh})/\mathrm{NP}^n_{\#wh}$
c. ghiseach; **girl**; N^n
d. tuische; **water**; N^n
e. $\lambda\sigma_1\lambda\varphi_1.\mathsf{aL}\circ\sigma_1(\varphi_1)$; $\lambda P.P$; $(\mathrm{S}'^{n+1}\upharpoonright\mathrm{NP}^{n+1}_{+wh})\upharpoonright(\mathrm{S}^n\upharpoonright\mathrm{NP}^n_{+wh})$

4. We do not, however, follow Pogodalla and Pompigne (2012) in encoding island constraints as syntactic constraints. See chapter 10 for our view on island constraints, where, following many authors in the recent literature, we take (most of) island constraints to follow from extragrammatical factors such as processing difficulty or prosodic/pragmatic incongruence. However, even assuming island constraints to be essentially extragrammatical, syntactic factors that are relevant for the formulations of such extragrammatical principles need to be captured by some kind of interface conditions making reference to essentially syntactic properties such as levels of clausal embedding. Thus, although we do not literally adopt Pogodalla and Pompigne's (2012) syntactic encoding of (a subtype of) island constraints, the formal machinery that they propose will very likely turn out to be essential in formulating explicitly the relevant interface conditions mediating syntax and other components of grammar (including the online processing component).

f. goN; $\lambda p.p$; S'^{n+1}/S^n

g. $\lambda\sigma_2\lambda\varphi_2.\varphi_2 \circ \sigma_2(\boldsymbol{\varepsilon})$; $\lambda P\lambda Q\lambda y.Q(y) \wedge P(y)$; $(N^n\backslash N^n)\upharpoonright(S'^n\upharpoonright NP^n_{+wh})$

(349g) is the null relativizer that combines the modifying property sign with the nominal predicates (349c) and (349d). With this lexicon, we obtain the following derivation for (350).

(350) An ghirseach aL dól an t-uisce
'the girl that drank the water'

(351)

$$\dfrac{\text{dól};\ \lambda x\lambda y.\mathbf{drink}(y)(x);\ (S^1/NP^1_{\#wh})/NP^1_{\#wh} \qquad \begin{bmatrix}\varphi_0;\\ z;\\ NP^1_{+wh}\end{bmatrix}^0}{\text{dól}\circ\varphi_0;\ \lambda y.\mathbf{drink}(y)(z);\ S^1/NP^1_{\#wh}}$$

$$\dfrac{\text{dól}\circ\varphi_0;\ \lambda y.\mathbf{drink}(y)(z);\ S^1/NP^1_{\#wh} \qquad \begin{matrix}\vdots\\ \text{an}\circ\text{tuische};\ \iota(\mathbf{water});\ NP^1_{-wh}\end{matrix}}{\text{dól}\circ\varphi_0\circ\text{an}\circ\text{tuische};\ \mathbf{drink}(\iota(\mathbf{water}))(z);\ S^1}$$

$$\dfrac{\text{dól}\circ\varphi_0\circ\text{an}\circ\text{tuische};\ \mathbf{drink}(\iota(\mathbf{water}))(z);\ S^1}{\lambda\varphi_0.\text{dól}\circ\varphi_0\circ\text{an}\circ\text{tuische};\ \lambda z.\mathbf{drink}(\iota(\mathbf{water}))(z);\ S^1\upharpoonright NP^1_{+wh}}\upharpoonright I^0$$

$$\dfrac{\lambda\varphi_0.\text{dól}\circ\varphi_0\circ\text{an}\circ\text{tuische};\ \lambda z.\mathbf{drink}(\iota(\mathbf{water}))(z);\ S^1\upharpoonright NP^1_{+wh} \qquad \lambda\sigma_1\lambda\varphi_1.\text{aL}\circ\sigma_1(\varphi_1);\ \lambda P.P;\ (S'^{n+1}\upharpoonright NP^{n+1}_{+wh})\upharpoonright(S^n\upharpoonright NP^n_{+wh})}{\lambda\varphi_1.\text{aL}\circ\text{dól}\circ\varphi_1\circ\text{an}\circ\text{tuische};\ \lambda z.\mathbf{drink}(\iota(\mathbf{water}))(z);\ S'^2\upharpoonright NP^2_{+wh}}$$

$$\dfrac{\lambda\sigma_2\lambda\varphi_2.\varphi_2\circ\sigma_2(\boldsymbol{\varepsilon});\ \lambda P\lambda Q\lambda y.Q(y)\wedge P(y);\ (N^2\backslash N^2)\upharpoonright(S'^2\upharpoonright NP^2_{\#wh}) \qquad \lambda\varphi_1.\text{aL}\circ\text{dól}\circ\varphi_1\circ\text{an}\circ\text{tuische};\ \lambda z.\mathbf{drink}(\iota(\mathbf{water}))(z);\ S'^2\upharpoonright NP^2_{+wh}}{\lambda\varphi_2.\varphi_2\circ\text{aL}\circ\text{dól}\circ\boldsymbol{\varepsilon}\circ\text{an}\circ\text{tuische};\ \lambda Q\lambda y.Q(y)\wedge\mathbf{drink}(\iota(\mathbf{water}))(y);\ N^2\backslash N^2}$$

$$\dfrac{\lambda\varphi_2.\varphi_2\circ\text{aL}\circ\text{dól}\circ\boldsymbol{\varepsilon}\circ\text{an}\circ\text{tuische};\ \lambda Q\lambda y.Q(y)\wedge\mathbf{drink}(\iota(\mathbf{water}))(y);\ N^2\backslash N^2 \qquad \text{ghiseach};\ \mathbf{girl};\ N^2}{\text{ghiseach}\circ\text{aL}\circ\text{dól}\circ\boldsymbol{\varepsilon}\circ\text{an}\circ\text{tuische};\ \lambda y.\mathbf{girl}(y)\wedge\mathbf{drink}(\iota(\mathbf{water}))(y);\ N^2}$$

$$\dfrac{\text{an};\ \lambda P.\iota(P);\ NP^2_{-wh}/N^2 \qquad \text{ghiseach}\circ\text{aL}\circ\text{dól}\circ\boldsymbol{\varepsilon}\circ\text{an}\circ\text{tuische};\ \lambda y.\mathbf{girl}(y)\wedge\mathbf{drink}(\iota(\mathbf{water}))(y);\ N^2}{\text{an}\circ\text{ghiseach}\circ\text{aL}\circ\text{dól}\circ\boldsymbol{\varepsilon}\circ\text{an}\circ\text{tuische};\ \iota(\mathbf{girl}(y)\wedge\mathbf{drink}(\iota(\mathbf{water}))(y));\ NP^2_{-wh}}$$

The key point to note here is that the empty relativizer requires both the body of the relative clause and the gap NP to carry the same clause-level index as the modified noun. This requirement can be satisfied by using *aL*, which explicitly passes the gap NP from the lower clause to the higher clause by incrementing the clause-level index of the gap NP by one.

Note in particular that using *goN* instead of *aL* here results in a failed derivation:

(352)

$$\dfrac{\dfrac{\dfrac{\dfrac{\vdots}{\textsf{dól} \circ \varphi_0 \circ \textsf{an} \circ \textsf{tuische};\ \mathbf{drink}(\iota(\mathbf{water}))(z);\ S^1} \quad \textsf{goN};\ \lambda p.p;\ S'^{n+1}/S^n}{\textsf{goN} \circ \textsf{dól} \circ \varphi_0 \circ \textsf{an} \circ \textsf{tuische};\ \mathbf{drink}(\iota(\mathbf{water}))(z);\ S'^2}\ {\upharpoonright}I^0}{\lambda\varphi_0.\textsf{goN} \circ \textsf{dól} \circ \varphi_0 \circ \textsf{an} \circ \textsf{tuische};\ \lambda z.\mathbf{drink}(\iota(\mathbf{water}))(z);\ S'^2{\upharpoonright}NP^1_{+wh}} \quad \begin{array}{l}\lambda\sigma_2\lambda\varphi_2.\varphi_2 \circ \sigma_2(\boldsymbol{\varepsilon});\\ \lambda P\lambda Q\lambda y.Q(y) \wedge P(y);\\ (N^n\backslash N^n){\upharpoonright}(S'^n{\upharpoonright}NP^n_{\#wh})\end{array}}{\mathbf{FAIL}}$$

goN increments the index of the S it selects, but it does not increment the index of the gap NP. The consequence of this is that when the NP_{+wh} hypothesis is withdrawn, the resultant expression has a syntactic type that is not compatible with the index requirement imposed by the relativizer.

This analysis entails the multiple appearances of *aL*, and only *aL*, along an extended filler-gap linkage exactly where we find them. Consider the following example from McCloskey (1979):

(353) an fear **aL** dúirt me **aL** shil me **aL** bheadh __ ann
'The man that I said I thought would be there.'

Assuming the following lexical entries for verbs that take clausal complements in (354), the derivation for (350) now proceeds as in (355)–(358).

(354) a. dúirt; $\lambda x\lambda p.\mathbf{said}(p)(x);\ S^n/S'^n/NP^n_{\#wh}$
b. shíl; $\lambda x\lambda p.\mathbf{think}(p)(x);\ S^n/S'^n/NP^n_{\#wh}$

In the first stage of the derivation shown in (355), we obtain the innermost appearance of *aL*, corresponding in phrase structure–based approaches to the smallest clause actually "housing" the extraction site:

(355)

$$\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\begin{array}{l}\textsf{bheadh};\\ \lambda x\lambda y.\mathbf{exist}(y)(x);\\ S^n/PP^n/NP^n_{\#wh}\end{array} \quad \begin{bmatrix}\varphi_0;\\ w;\\ NP^1_{+wh}\end{bmatrix}^0}{\textsf{bheadh} \circ \varphi_0;\ \lambda y.\mathbf{exist}(y)(w);\ S^1/PP^1} \quad \begin{array}{l}\textsf{ann};\\ \mathbf{there};\\ PP^1\end{array}}{\textsf{bheadh} \circ \varphi_0 \circ \textsf{ann};\ \mathbf{exist}(\mathbf{there})(w);\ S^1}\ {\upharpoonright}I^0}{\begin{array}{l}\lambda\varphi_0.\textsf{bheadh} \circ \varphi_0 \circ \textsf{ann};\\ \lambda w.\mathbf{exist}(\mathbf{there})(w);\ S^1{\upharpoonright}NP^1_{+wh}\end{array}} \quad \begin{array}{l}\lambda\sigma_1\lambda\varphi_1.\textsf{aL} \circ \sigma_1(\varphi_1);\\ \lambda P.P;\\ (S'^{n+1}{\upharpoonright}NP^{n+1}_{+wh}){\upharpoonright}(S^n{\upharpoonright}NP^n_{+wh})\end{array}}{\lambda\varphi_1.\textsf{aL} \circ \textsf{bheadh} \circ \varphi_1 \circ \textsf{ann};\ \lambda w.\mathbf{exist}(\mathbf{there})(w);\ S'^2{\upharpoonright}NP^2_{+wh}} \quad \begin{bmatrix}\varphi_2;\\ v;\\ NP^2_{+wh}\end{bmatrix}^2}{\textsf{aL} \circ \textsf{bheadh} \circ \varphi_1 \circ \textsf{ann};\ \mathbf{exist}(\mathbf{there})(v);\ S'^2}$$

At this point, at the penultimate proof step, *aL* has been composed into the clause representing the initial extraction domain. To complete this part of the derivation, we then saturate the resulting $\mathrm{S}\restriction\mathrm{NP}$ sign to make it eligible for selection by *shíl*.

(356)

$$\dfrac{\dfrac{\dfrac{\dfrac{\begin{array}{l}\text{shíl};\\ \lambda x\lambda p.\mathbf{think}(p)(x);\\ \mathrm{S}^n/\mathrm{S}'^n/\mathrm{NP}^n_{\#wh}\end{array} \quad \begin{array}{l}\text{me};\\ \mathbf{1st};\\ \mathrm{NP}^2_{-wh}\end{array}}{\begin{array}{l}\text{shíl}\circ\text{me};\\ \lambda p.\mathbf{think}(p)(\mathbf{1st});\mathrm{S}^2/\mathrm{S}'^2\end{array}} \quad \begin{array}{l}\vdots\\ \text{aL}\circ\text{bheadh}\circ\\ \quad\varphi_2\circ\text{ann};\\ \mathbf{exist}(\mathbf{there})(v);\\ \mathrm{S}'^2\end{array}}{\begin{array}{l}\text{shíl}\circ\text{me}\circ\text{aL}\circ\text{bheadh}\circ\varphi_2\circ\text{ann};\\ \mathbf{think}(\mathbf{exist}(\mathbf{there})(v))(\mathbf{1st});\mathrm{S}^2\end{array}}\restriction\mathrm{I}^2}{\begin{array}{l}\lambda\varphi_2.\text{shíl}\circ\text{me}\circ\text{aL}\circ\text{bheadh}\circ\varphi_2\circ\text{ann};\\ \lambda v.\mathbf{think}(\mathbf{exist}(\mathbf{there})(v))(\mathbf{1st});\mathrm{S}^2\restriction\mathrm{NP}^2_{+wh}\end{array}} \quad \begin{array}{l}\lambda\sigma_2\lambda\varphi_3.\text{aL}\circ\sigma_2(\varphi_3);\\ \lambda P.P;\\ (\mathrm{S}'^{n+1}\restriction\mathrm{NP}^{n+1}_{+wh})\restriction(\mathrm{S}^n\restriction\mathrm{NP}^n_{+wh})\end{array}}{\begin{array}{l}\lambda\varphi_3.\text{aL}\circ\text{shíl}\circ\text{me}\circ\text{aL}\circ\text{bheadh}\circ\varphi_3\circ\text{ann};\\ \lambda v.\mathbf{think}(\mathbf{exist}(\mathbf{there})(v))(\mathbf{1st});\mathrm{S}^3\restriction\mathrm{NP}^3_{+wh}\end{array}} \quad \begin{bmatrix}\varphi_4;\\ z;\\ \mathrm{NP}^3_{+wh}\end{bmatrix}^4}{\text{aL}\circ\text{shíl}\circ\text{me}\circ\text{aL}\circ\text{bheadh}\circ\varphi_4\circ\text{ann};\ \mathbf{think}(\mathbf{exist}(\mathbf{there})(z))(\mathbf{1st});\ \mathrm{S}'^3}$$

That is, once *shíl* has composed into the derivation, we bind the new φ_2 variable—thus in effect renewing the abstraction corresponding to the original gap in the first step of the proof. A final iteration of the same proof sequence—in abstract terms, λ-bind a variable, saturate the associated functor, compose the result with a new selector, and then rebind the variable—will yield the final form of the relative clause:

(357)

$$\dfrac{\dfrac{\dfrac{\begin{array}{l}\text{dúirt};\\ \lambda x\lambda p.\mathbf{said}(p)(x);\\ \mathrm{S}^n/\mathrm{S}'^n/\mathrm{NP}^n_{\#wh}\end{array} \quad \begin{array}{l}\text{me};\\ \mathbf{1st};\mathrm{NP}^3_{-wh}\end{array}}{\text{dúirt}\circ\text{me};\ \lambda p.\mathbf{said}(p)(\mathbf{1st});\ \mathrm{S}^n/\mathrm{S}'^n} \quad \begin{array}{l}\vdots\\ \text{aL}\circ\text{shíl}\circ\text{me}\circ\text{aL}\circ\text{bheadh}\circ\varphi_4\circ\text{ann};\\ \mathbf{think}(\mathbf{exist}(\mathbf{there})(z))(\mathbf{1st});\mathrm{S}'^3\end{array}}{\begin{array}{l}\text{dúirt}\circ\text{me}\circ\text{aL}\circ\text{shíl}\circ\text{me}\circ\text{aL}\circ\text{bheadh}\circ\varphi_4\circ\text{ann};\\ \mathbf{said}(\mathbf{think}(\mathbf{exist}(\mathbf{there})(z))(\mathbf{1st}))(\mathbf{1st});\mathrm{S}^3\end{array}}\restriction\mathrm{I}^4}{\begin{array}{l}\lambda\varphi_4.\text{dúirt}\circ\text{me}\circ\text{aL}\circ\text{shíl}\circ\text{me}\circ\text{aL}\circ\text{bheadh}\circ\varphi_4\circ\text{ann};\\ \lambda z.\mathbf{said}(\mathbf{think}(\mathbf{exist}(\mathbf{there})(z))(\mathbf{1st}))(\mathbf{1st});\mathrm{S}^3\restriction\mathrm{NP}^3_{+wh}\end{array}}$$

The proof concludes with the linkage among the relative clause, the null relativizer, and the "head" noun of the NP:

(358)

$$\vdots$$

$\lambda\varphi_4.$dúirt ∘ me ∘ aL ∘ shíl ∘ me ∘ aL ∘ bheadh ∘ φ_4 ∘ ann; $\lambda z.\mathbf{said}(\mathbf{think}(\mathbf{exist}(\mathbf{there})(z))(\mathbf{1st}))(\mathbf{1st})$; $S^3{\upharpoonright}NP^3_{+wh}$

$\lambda\sigma_1\lambda\varphi_1.$aL ∘ $\sigma_1(\varphi_1)$; $\lambda P.P$; $(S'^{n+1}{\upharpoonright}NP^{n+1}_{+wh}){\upharpoonright}(S^n{\upharpoonright}NP^n_{+wh})$

$\lambda\varphi_4.$aL ∘ dúirt ∘ me ∘ aL ∘ shíl ∘ me ∘ aL ∘ bheadh ∘ φ_4 ∘ ann; $\lambda z.\mathbf{said}(\mathbf{think}(\mathbf{exist}(\mathbf{there})(z))(\mathbf{1st}))(\mathbf{1st})$; $S^4{\upharpoonright}NP^4_{+wh}$

$\lambda\sigma_2\lambda\varphi_5.\varphi_5 \circ \sigma_2(\boldsymbol{\varepsilon})$; $\lambda P\lambda Q\lambda y.Q(y)\wedge P(y)$; $(N^n\backslash N^n){\upharpoonright}(S'^n{\upharpoonright}NP^n_{+wh})$

$\lambda\varphi_5.\varphi_5$ ∘ aL ∘ dúirt ∘ me ∘ aL ∘ shíl ∘ me ∘ aL ∘ bheadh ∘ $\boldsymbol{\varepsilon}$ ∘ ann; $\lambda Q\lambda y.Q(y)\wedge\mathbf{said}(\mathbf{think}(\mathbf{exist}(\mathbf{there})(y))(\mathbf{1st}))(\mathbf{1st})$; $N^4\backslash N^4$

fear; **man**; N^4

---

fear ∘ aL ∘ dúirt ∘ me ∘ aL ∘ shíl ∘ me ∘ aL ∘ bheadh ∘ $\boldsymbol{\varepsilon}$ ∘ ann; $\lambda y.\mathbf{man}(y)\wedge\mathbf{said}(\mathbf{think}(\mathbf{exist}(\mathbf{there})(y))(\mathbf{1st}))(\mathbf{1st})$; $N^4\backslash N^4$

an; $\lambda P.\iota(P)$; NP^4_{-wh}/N^4

---

an ∘ fear ∘ aL ∘ dúirt ∘ me ∘ aL ∘ shíl ∘ me ∘ aL ∘ bheadh ∘ $\boldsymbol{\varepsilon}$ ∘ ann; $\iota(\lambda y.\mathbf{man}(y)\wedge\mathbf{said}(\mathbf{think}(\mathbf{exist}(\mathbf{there})(y))(\mathbf{1st}))(\mathbf{1st}))$; NP^4_{-wh}

Note further that examples such as those in (359) are correctly predicted to be impossible on this analysis:

(359) a. *an fear **aL** duiirt me **aL** shil me **goN** bheadh __ ann
'The man that I said I thought would be there.'

b. *an fear **goN** duiirt me **aL** shil me **aL** bheadh __ ann
'The man that I said I thought would be there.'

In the case of (359a), *goN* will combine with a clause containing a hypothesized NP^1_{+wh}, shifting the index of that clause to 2, and the *shíl* clause will maintain this index on the resulting S′, as will *dúirt*. In order to combine with *aL*, however, the original NP^1_{+wh} hypothesis at the "bottom" will have to be withdrawn from the *dúirt* clause, yielding an argument of the form $S^2{\upharpoonright}NP^1_{+wh}$—an unacceptable argument type for *aL*, which requires the same index on both argument and yield. More generally, anytime *goN* appears between a filler and a gap, it will have been necessary to saturate the functor corresponding to the gap with an NP (or other filler type) at the same index level as its selector, since *goN* requires a complete S to combine with. But any *aL* higher up requires an S↾NP category where the argument and the result share the same index value. This will be impossible, since once the hypothesized NP that saturated the *goN* clause is withdrawn, the result—due to the index effect of *goN*—will be a constituent of the form $S^n{\upharpoonright}NP^m$ with $n > m$. If the hypothesis is withdrawn at the very top of the dependency, as in (359b), then the result will fail for essentially the same reason, since not just *aL* but the relativizer and all *wh* terms require arguments of the form $S^n{\upharpoonright}NP^n$.

Thus *goN* is blocked from any appearance between a filler and a "gap," and it is also obvious that *aL* will be blocked in any context where there is no extraction dependency (note that the "missing" NP in the lexical specification for *aL* is specifically marked as +*wh*). The complementarity between *goN* and *aL*, and the mandatory appearance of *aL* at every point along a filler gap pathway, thus follows directly. For if *aL* is missing at some point, then the string in question will be typed $\mathrm{S}^n{\upharpoonright}\mathrm{NP}^n$, rather than the $\mathrm{S}'^n{\upharpoonright}\mathrm{NP}^n$ required of arguments to Irish predicates seeking clausal complements. Hence *aL* is obligatory in each clausal sign along the extraction pathway, just as *goN* is obligatory for clauses in nonextraction constructions.[5]

7.2 Linearity of the Underlying Logic and the Treatment of Parasitic Gaps

Multiple-gap phenomena (of which parasitic gaps are an instance) perhaps pose a somewhat more serious issue for the treatment of extraction in terms of the linear implication connective $\upharpoonright$ as the key device for mediating the linkage between the filler and the gap. While extraction pathway marking can be modeled by encoding the level of clause embedding via a syntactic feature (a mechanism which can potentially be retooled for other purposes along lines briefly discussed above), the fact that at least in some languages (including English) a single filler can correspond to multiple gaps simultaneously, as in the following parasitic gap examples in English, seems to be directly at odds with the assumption that the underlying logic for the combinatoric component of natural language grammar is resource-sensitive linear logic.

(360) a. This is the article that John filed __ without reading __.
b. Peter is a guy who even the best friends of __ tend to avoid hanging around with __.

Unlike the possibilities available in classical logic, in which $A, A \to (A \to B) \vdash B$ is a valid theorem (by using the minor premise A twice in Modus Ponens), in linear logic, the corresponding inference $A, A \multimap (A \multimap B) \vdash B$ is *not* a theorem since reuse of a resource is not an option. But if we model the extraction phenomena via hypothetical reasoning, along the lines of Muskens (2003), the derivation for multiple-gap sentences would seem to necessarily correspond to the unavailable theorem in linear logic.

Does this mean, after all, that the implication connective mediating long-distance dependencies in English (and other languages that allow for multiple-gap constructions)

5. Readers familiar with the syntax-semantics interface presented in Gazdar et al. (1985) will find our analysis of "extraction pathway marking" as the clause-by-clause saturation and rebinding of the variable associated with the prosodic gap site rather familiar: it is simply the TLCG version of the way Gazdar et al. treated filler-gap interpretation as the semantic correspondent of the SLASH feature's locally mediated distribution through all clause levels separating fillers from their matching gap sites.

is *not* linear? Some authors have indeed drawn this conclusion and have proposed analyses which explicitly reject linearity, at least in contexts in which multiple gaps are licensed in extraction constructions. The most recent and the most detailed proposal embodying this idea can be found in Morrill (2017). Here, we opt for a less radical departure from the assumption of linearity, not because we find the extensive revision of the underlying logic undertaken by Morrill (2017) implausible, but merely because we wish to remain agnostic about this potentially quite deep and fundamental question that has direct ramifications to the architecture of the syntax-semantics interface. For this reason, we sketch below a lexical operator–based analysis of parasitic gaps in English which retains the linearity of the underlying logic and introduces resource duplication via an empty operator that is associated with a nonlinear term as its semantic translation. We then formulate our account of the Kennedy ellipsis/extraction paradigm in the next chapter on the basis of this analysis of parasitic gaps.[6]

We can overcome the technical difficulty noted above by positing a lexical operator that introduces multiple tokens of a variable which are "pre-bound," along the lines of (361):[7]

(361) $\lambda\sigma_1\lambda\sigma_2\lambda\varphi.\sigma_2(\varphi)\circ\sigma_1(\varphi)$;
$\lambda R\lambda g\lambda x.R(x)(g(x))$; $(\mathrm{S}\upharpoonright\mathrm{NP}_{+wh})\upharpoonright(\mathrm{NP}_{-wh}\upharpoonright\mathrm{NP}_{+wh})\upharpoonright((\mathrm{NP}_{-wh}\backslash\mathrm{S})\upharpoonright\mathrm{NP}_{+wh})$

6. As will become clear in the next chapter, the only crucial assumption for our analysis of the Kennedy paradigm is that the "existence" of a gap within an "elided" VP is marked in the syntactic category of the latter. In order to see whether our analysis carries over to other TLCG approaches, detailed assumptions about how exactly the analyses of ellipsis and extraction interact with one another in each approach needs to be spelled out. This is of course not a trivial task, but we believe that a translation of our analysis to other TLCG approaches (including Morrill [2017]) should be mostly straightforward, since the encoding of extracted elements in the category of linguistic signs is a property that is for the most part orthogonal to the question of whether the implicational connective employed is linear or not. In any event, regardless of what conclusion one draws on the linearity issue, we take the Kennedy paradigm to be one of the empirical criteria for judging the adequacy of an analysis of extraction phenomena (in particular, the licensing of multiple gaps) in English.

7. This operator, on its own, licenses parasitic gaps in the subject position only. In order to license parasitic gaps in adjunct clauses, a second, separate operator is needed. One might object to this type of analysis on the ground that it fails to capture parasitic gaps in the two environments via a uniform mechanism. However, it is questionable whether a uniform analysis is desirable to begin with. Culicover (1999, 179–181) notes that "there appear to be many languages that have *without*-parasitic gaps and parasitic gaps in adjuncts, but lack parasitic gaps in subjects," giving Spanish and German as two such languages. We therefore take the need for separate operators for adjunct and subject parasitic gaps to be well supported empirically. Questions, however, remain regarding the much larger issue of how to deal with a wider range of multiple-gap phenomena in Hybrid TLCG. We can find in English a variety of other multiple-gap/single-filler constructions, including "parasitic" gaps in adjuncts, ATB extraction in coordination, the "symbiotic" gaps pointed out in Levine and Sag (2003), and cases such as (i), where both gaps are in extractable positions within a single VP:

(i) Which people$_i$ did you show pictures of $\_\__i$ to $\_\__i$?

We leave it for future work to extend our analysis to such cases.

This operator takes a VP missing an NP and an NP missing an NP and in effect fuses the missing NP argument variables, and likewise for the prosodic variables. The result, supplied as an argument to the *wh* operator, yields a sign with the empty string in two separate positions, corresponding to two tokens of a single variable in the semantic component of the sign, bound by a single abstraction operator.[8] The action of the gap-multiplying operator (361) is illustrated in (362):

(362)

$$\lambda\sigma_1\lambda\sigma_2\lambda\varphi.\ \sigma_2(\varphi)\circ\sigma_1(\varphi);\ \lambda R\lambda g\lambda x.R(x)(g(x));\ (\mathrm{S}{\restriction}\mathrm{NP}_{+wh}){\restriction}(\mathrm{NP}_{-wh}{\restriction}\mathrm{NP}_{+wh}){\restriction}((\mathrm{NP}_{-wh}\backslash\mathrm{S}){\restriction}\mathrm{NP}_{+wh})$$

$$\vdots$$

$$\lambda\varphi.\mathsf{admire}\circ\varphi;\ \mathbf{admire};\ (\mathrm{NP}_{-wh}\backslash\mathrm{S}){\restriction}\mathrm{NP}_{+wh}$$

$$\overline{\lambda\sigma_2\lambda\varphi.\sigma_2(\varphi)\circ\mathsf{admire}\circ\varphi;\ \lambda g\lambda x.\mathbf{admire}(x)(g(x));\ (\mathrm{S}{\restriction}\mathrm{NP}_{+wh}){\restriction}(\mathrm{NP}_{-wh}{\restriction}\mathrm{NP}_{+wh})}$$

$$\vdots$$

$$\lambda\varphi.\mathsf{the}\circ\mathsf{close}\circ\mathsf{friends}\circ\mathsf{of}\circ\varphi;\ \lambda x.\iota(\lambda y.\mathbf{close\text{-}fr}(x)(y));\ \mathrm{NP}_{-wh}{\restriction}\mathrm{NP}_{+wh}$$

$$\overline{\lambda\varphi.\mathsf{the}\circ\mathsf{close}\circ\mathsf{friends}\circ\mathsf{of}\circ\varphi\circ\mathsf{admire}\circ\varphi;\ \lambda x.\mathbf{admire}(x)(\iota(\lambda y.\mathbf{close\text{-}fr}(x)(y)));\ \mathrm{S}{\restriction}\mathrm{NP}_{+wh}}$$

$$\lambda\sigma.\mathsf{who}\circ\sigma(\boldsymbol{\varepsilon});\ \lambda P\lambda Q\lambda x.\ Q(x)\wedge P(x);\ (\mathrm{N}\backslash\mathrm{N}){\restriction}(\mathrm{S}{\restriction}\mathrm{NP}_{+wh})$$

$$\overline{\mathsf{who}\circ\mathsf{the}\circ\mathsf{close}\circ\mathsf{friends}\circ\mathsf{of}\circ\mathsf{admire};\ \lambda Q\lambda x.Q(x)\wedge\mathbf{admire}(x)(\iota(\lambda y.\mathbf{close\text{-}fr}(x)(y)));\ \mathrm{N}\backslash\mathrm{N}}$$

8. A similar strategy for dealing with nonlinearity in natural language is pursued in the cross-categorial analysis of coordination via "generalized conjunction" by Partee and Rooth (1983) and analyses which build the distribution of reflexive pronouns into the semantics of such pronouns along the lines of Bach and Partee (1980).