The P-stranding generalization does not require covert structure

Yusuke Kubota
University of Tsukuba

Robert Levine
Ohio State University

September 7, 2017

The P-stranding generalization Merchant (2001) contains an argument for covert structure in sluicing based on what has come to be known as the P-stranding generalization:

(1) A language *L* will allow preposition stranding under sluicing iff *L* allows preposition stranding under regular *wh*-movement. (Merchant 2001, 92)

The empirical basis for this claim is a set of data from eighteen languages: six Germanic languages which support P-stranding (both in the regular *wh*-movement context and in sluicing) and twelve from a wide variety of languages (including a variety of Indo-Iranian and Semitic languages, as well as Basque) which do not. On the basis of these facts and the descriptive generalization in (1), Merchant (2001, 107) argues that

> the usual mechanisms for case-assignment and determination of targets of *wh*-movement that operate in a given language to regulate the shapes of *wh*-phrases in non-elliptical questions operate in identical ways under sluicing as well. All of these facts strongly suggest that *wh*-movement of the usual sort has taken place, displacing an IP-internal *wh*-phrase to SpecCP ... similar considerations suggest a movement approach to a variety of parallel ... form-identify effects in stripping, comparatives, fragment answers, [and] the remnants of gapping, which often show case and P-stranding dependencies like their sluicing cousins

By the same reasoning, the putative generalization should be observable in examples like the following, where (2a) is standardly analyzed as extraction out of elided VPs (but see Kubota and Levine 2017a) and (2b) is a case of pseudogapping:

(2) a. I know whom John argued with, but I don't know whom Mary did.
b. I can deal with Mary more easily than I can Sue.

In (2a) the *wh* filler has moved to the left, leaving behind a VP with a trace in it (*argued with* __), which is subsequently deleted on Merchant's analysis. In a language that allows extraction out of VP-ellipsis sites but which has a ban on P-stranding, data such as (2a) should never exist, and likewise for analogues of (2b).[1]

[1]Matters may be a bit more complex in the case of pseudogapping, though. On some analyses, remnant movement in pseudogapping is to the right, leaving open the possibility that in a language with a P-stranding prohibition on leftward movement only, something like (2b) could be legal.

Apparent exceptions to the P-stranding generalization exist. For example, Sag and Nykiel (2011) argue that Polish counterexemplifies this generalization; Merchant (2013) himself refers to the literature on several such cases and suggests the possibility that in these cases 'the P-less 'sluices' in fact derive from a copular or reduced cleft-like source', a possibility which he refers to as pseudosluicing. Recent work however challenges this alternative derivation of apparent exception: Nykiel (2013) offers a suite of experimental psycholinguistic tests whose results suggest that in general the acceptability of sluicing with a lone preposition remnant is unrelated to the acceptability of *wh* extraction of NPs from PPs. The issue thus seems sufficiently vexed that no firm conclusions are possible at this point. In the discussion below, we will simply assume that Merchant's P-stranding generalization is correct, and demonstrate that even if it is, it has no implications for the existence of hierarchical structure at the ellipsis site.

Marking 'extracted' NPs via a syntactic feature We begin by outlining how the pattern described as extraction in phrase-structure-based approaches is accounted for in Hybrid Type-Logical Categorial Grammar (H-TLCG), the core proof theory of which is given in (3):[2]

(3)

Connective	Introduction	Elimination
$/$	$\dfrac{\begin{array}{c}[\varphi; x; A]^n \\ \vdots \quad \vdots \\ \mathsf{b} \circ \varphi; \mathcal{F}; B\end{array}}{\mathsf{b}; \lambda x.\mathcal{F}; B/A}\, /\,\mathrm{I}^n$	$\dfrac{\mathsf{a}; \mathcal{F}; A/B \quad \mathsf{b}; \mathcal{G}; B}{\mathsf{a} \circ \mathsf{b}; \mathcal{F}(\mathcal{G}); A}\, /\mathrm{E}$
$\backslash$	$\dfrac{\begin{array}{c}[\varphi; x; A]^n \\ \vdots \quad \vdots \\ \varphi \circ \mathsf{b}; \mathcal{F}; B\end{array}}{\mathsf{b}; \lambda x.\mathcal{F}; A\backslash B}\, \backslash\mathrm{I}^n$	$\dfrac{\mathsf{b}; \mathcal{G}; B \quad \mathsf{a}; \mathcal{F}; B\backslash A}{\mathsf{b} \circ \mathsf{a}; \mathcal{F}(\mathcal{G}); A}\, \backslash\mathrm{E}$
$\upharpoonright$	$\dfrac{\begin{array}{c}[\varphi; x; A]^n \\ \vdots \quad \vdots \\ \mathsf{b}; \mathcal{F}; B\end{array}}{\lambda\varphi.\mathsf{b}; \lambda x.\mathcal{F}; B\upharpoonright A}\, \upharpoonright\mathrm{I}^n$	$\dfrac{\mathsf{a}; \mathcal{F}; A\upharpoonright B \quad \mathsf{b}; \mathcal{G}; B}{\mathsf{a}(\mathsf{b}); \mathcal{F}(\mathcal{G}); A}\, \upharpoonright\mathrm{E}$

We assume the analysis of extraction sketched in Muskens (2003), based on lexical operators of the kind exhibited in (4) (see Kubota and Levine (2017b,a) for a compact exposition of this analysis of extraction):

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

Here ϵ the empty string. In the semantic term, **obj** identifies the set of objects, and **wh** is a shorthand for some appropriate semantics for *wh* questions.[3]

[2]For an exposition of H-TLCG, see Kubota and Levine (2014). For an application of H-TLCG to a broad range of empirical problems, see references cited therein. Kubota and Levine (2017b) and Kubota and Levine (2017a) contain our main results on the analysis of ellipsis phenomena so far.

[3]One can, for example, adopt the Karttunen semantics for questions (Karttunen 1977), which identifies

The notation NP_{+wh} requires somewhat further comment. The extraction operator in (4) maps an S missing an NP (e.g., *did John like* __) to a question (*Who did John like?*). As we illustrate below, the derivation of such sentences, with the gap site arbitrarily distant from the filler, involves no intermediate structural representations. Yet there are genuine locality effects which must be taken into account in any analysis of extraction phenomena—what Zaenen (1983) characterizes as 'syntactic binding domain effects', with filler-gap pathways distinguished by unique morphosyntactic or phonological signatures in some languages (such as Icelandic). Our ongoing research suggests that these locality effects can be elegantly accounted for without any appeal to hierarchical configuration if we assume that arguments which are bindable by extraction operators such as (4) are distinguished from other NPs. The binary *wh* feature is introduced for this purpose.[4] We assume that NPs unspecified for the *wh* feature is a join of these two more specific types, i.e., $NP \equiv NP_{+wh} \vee NP_{-wh}$ (see Morrill (1994), Bayer and Johnson (1995) and Bayer (1996) for how features and feature underspecification in unification-based grammars can be modelled by the meet and join connectives in categorial grammar).

This modest elaboration of the type system has an apparently significant technical consequence (which however turns out to be a non-issue). Consider a very simple sentence such as (5):

(5) I wonder what John ate.

In order to supply a variable whose abstraction can feed the extraction operator (4), 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 VP/NP, where the NP argument is unspecified for *wh* value. This seems to create a problem at the very outset of the derivation for (5):

(6)
$$\frac{\textsf{ate}; \mathbf{eat}; \mathrm{VP/NP} \quad [\varphi_0; x; \mathrm{NP}_{+wh}]^1}{\mathbf{??}}\,/\mathrm{E}$$

The inference rules in (3) explicitly require identity in the argument type of a functor and the type of the sign it combines with. Since NP ($\equiv NP_{+wh} \vee NP_{-wh}$) and NP_{+wh} are not identical, it appears that /E cannot apply, leaving us with no way to derive (5). But there is no real difficulty: disjunctive categories have their own inference rules, which are independently motivated empirically in the analysis of feature neutralization effects, as discussed in detail in Bayer (1996). We reproduce here the ∨ Introduction rules from Bayer (1996):[5]

(7)
$$\frac{\textsf{a}; \Gamma; X}{\textsf{a}; \Gamma; X \vee Y} \qquad \frac{\textsf{b}; \Gamma; Y}{\textsf{b}; \Gamma; X \vee Y}$$

the meaning of a *wh* question with the set of propositions that count as true answers to that question. The particular choice of the semantics is irrelevant to the issues we discuss below.

[4]Note that Steedman (1996) employs essentially the same technique (his ±ANT feature), for a somewhat different purpose.

[5]Here we follow Bayer's assumption that terms inhabiting conjoined and disjoined types cannot differ in their semantics, regardless of which subtype they belong to. Such sign, 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. The Introduction rule in (7) therefore has no effect on the meaning of the expression belonging to the resulting complex type.

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

(8)
$$\dfrac{\dfrac{k;\ \mathrm{VP}/(\mathrm{NP}_{+wh} \vee \mathrm{NP}_{-wh}) \qquad \dfrac{[\varphi_0; \mathrm{NP}_{+wh}]^1}{\varphi_0;\ \mathrm{NP}_{+wh} \vee \mathrm{NP}_{-wh}}\,{\scriptstyle \vee\mathrm{I}}}{k \circ \varphi_0;\ \mathrm{VP}}\,{\scriptstyle /\mathrm{E}}}{k;\ \mathrm{VP}/\mathrm{NP}_{+wh}}\,{\scriptstyle /\mathrm{I}^1}$$

This proof can be trivially generalized schematically in the from 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 $\upharpoonright$). With this lemma abbreviated by a dashed line labelled '$\vee$Lemma', we have the following derivation:

(9)
$$\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\mathsf{ate};\ \mathbf{eat};\ \mathrm{VP}/\mathrm{NP}}{\mathsf{ate};\ \mathbf{eat};\ \mathrm{VP}/\mathrm{NP}_{+wh}}\,{\scriptstyle \vee\mathrm{Lemma}} \quad \begin{bmatrix}\varphi_0; \\ x; \mathrm{NP}_{+wh}\end{bmatrix}^1}{\mathsf{ate} \circ \varphi_0;\ \mathbf{eat}(x);\ \mathrm{VP}}\,{\scriptstyle /\mathrm{E}} \quad \begin{matrix}\mathsf{john}; \\ \mathbf{j}; \mathrm{NP}_{-wh}\end{matrix}}{\mathsf{john} \circ \mathsf{ate} \circ \varphi_0;\ \mathbf{eat}(x)(\mathbf{j});\ \mathrm{S}}\,{\scriptstyle \backslash\mathrm{E}}}{\lambda\varphi_0.\mathsf{john} \circ \mathsf{ate} \circ \varphi_0;\ \lambda x.\mathbf{eat}(x)(\mathbf{j});\ \mathrm{S}\upharpoonright\mathrm{NP}_{+wh}}\,{\scriptstyle \upharpoonright\mathrm{I}^1} \quad \begin{matrix}\lambda\sigma.\mathsf{what} \circ \sigma(\boldsymbol{\epsilon}); \\ \mathbf{wh}(\mathbf{obj}); \mathrm{Q}\upharpoonright(\mathrm{S}\upharpoonright\mathrm{NP}_{+wh})\end{matrix}}{\mathsf{what} \circ \mathsf{john} \circ \mathsf{ate} \circ \boldsymbol{\epsilon};\ \mathbf{wh}(\mathbf{obj})(\lambda x.\mathbf{eat}(x)(\mathbf{j}));\ \mathrm{Q}}\,{\scriptstyle \upharpoonright\mathrm{E}}}{}$$

English* We now consider a language English*, which is exactly like English except that preposition stranding is forbidden. To provide continuity with our treatment of ellipsis in Kubota and Levine (2017a,b), we first demonstrate how that treatment, with no additions or modifications, yields the P-stranding generalization for the so-called extraction out of elided VP cases like (2a) (for which we have defended an analysis (essentially) in terms of extraction of pseudogapping remnants). Assume that in English* we have only (10) as the lexical entry for the preposition *to*.

(10) $\mathsf{to};\ \lambda x.x;\ \mathrm{PP}_{to}/\mathrm{NP}_{-wh}$

This specification differs crucially from the lexical description of *to* in English in that the syntactic type of the latter will be PP/NP (see below), with both NP_{+wh} and NP_{-wh} as possible subtypes for the argument NP.

Given the lexical specifications for prepositions in English*, indirect questions such as *I wonder who John spoke to* cannot be formed. Since only an NP_{-wh} variable can be supplied to the object of *to*, a sentence missing an NP object of a preposition such as *John spoke to* __ can only be derived in type $\mathrm{S}\upharpoonright\mathrm{NP}_{-wh}$. But such a description fails to satisfy the fronted *wh*-word's argument description $\mathrm{S}\upharpoonright\mathrm{NP}_{+wh}$ (cf. (4)). And precisely the same will hold in the attempt to derive the corresponding elided VP. Consider (11):

(11) JOHN talked to BILL, but I don't know whom MARY did.

This example is acceptable in English (as long as the proper intonational and contextual cues are given), but would be ill-formed in English*. The antecedent upon which the ellipsis clause depends will have the following derivation:

(12)

$$
\dfrac{
\dfrac{
\dfrac{
\dfrac{
\dfrac{
\dfrac{
\left[\begin{array}{l}\varphi_0;\\ u; \mathrm{NP}_{-wh}\end{array}\right]^1
\quad
\begin{array}{l}\mathsf{to};\\ \lambda\varphi.\varphi;\\ \mathrm{PP}_{to}/\mathrm{NP}_{-wh}\end{array}
}{
\mathsf{to} \circ \varphi_0;\ u;\ \mathrm{PP}_{to}
}
\quad
\begin{array}{l}\mathsf{talked};\\ \mathbf{talk}; \mathrm{VP}/\mathrm{PP}_{to}\end{array}
}{
\mathsf{talked} \circ \mathsf{to} \circ \varphi_0;\ \mathbf{talk}(u);\ \mathrm{VP}
}
}{
\lambda\varphi_0.\mathsf{talked} \circ \mathsf{to} \circ \varphi_0;\ \lambda u.\mathbf{talk}(u);\ \mathrm{VP}{\upharpoonright}\mathrm{NP}_{-wh}
}
\quad
\left[\begin{array}{l}\varphi_1;\\ w; \mathrm{NP}_{-wh}\end{array}\right]^2
}{
\mathsf{talked} \circ \mathsf{to} \circ \varphi_1;\ \mathbf{talk}(w);\ \mathrm{VP}
}
\quad
\begin{array}{l}\mathsf{john};\\ \mathbf{j}; \mathrm{NP}\end{array}
}{
\dfrac{
\mathsf{john} \circ \mathsf{talked} \circ \mathsf{to} \circ \varphi_1;\ \mathbf{talk}(w)(\mathbf{j});\ \mathrm{S}
}{
\lambda\varphi_0.\mathsf{john} \circ \mathsf{talked} \circ \mathsf{to} \circ \varphi_0;\ \lambda w.\mathbf{talk}(w)(\mathbf{j});\ \mathrm{S}{\upharpoonright}\mathrm{NP}_{-wh}
}{\upharpoonright}\mathrm{I}^1
}
\quad
\begin{array}{l}\mathsf{bill};\\ \mathbf{b}; \mathrm{NP}_{-wh}\end{array}
}{
\mathsf{john} \circ \mathsf{talked} \circ \mathsf{to} \circ \mathsf{bill};\ \mathbf{talk}(\mathbf{b})(\mathbf{j});\ \mathrm{S}
}{\upharpoonright}\mathrm{E}
$$

And for the VPE auxiliary, following the analysis in Kubota and Levine (2017a), we obtain:

(13)

$$
\dfrac{
\dfrac{
\dfrac{
\begin{array}{l}\vdots \quad \vdots\\ \lambda\sigma\lambda\varphi.\mathsf{did} \circ \sigma(\varphi);\\ \lambda f\lambda x\lambda y.f(x)(y);\\ (\mathrm{VP}{\upharpoonright}\mathrm{NP}_{-wh}){\upharpoonright}(\mathrm{VP}{\upharpoonright}\mathrm{NP}_{-wh})\end{array}
\quad
\textcircled{1}\!\rightarrow
\begin{array}{l}\lambda\rho\lambda\varphi.\rho(\lambda\varphi_0.\varphi_0)(\varphi);\\ \lambda\mathscr{F}.\mathscr{F}(P);\\ (\mathrm{VP}{\upharpoonright}\mathrm{NP}_{-wh}){\upharpoonright}((\mathrm{VP}{\upharpoonright}\mathrm{NP}_{-wh}){\upharpoonright}(\mathrm{VP}{\upharpoonright}\mathrm{NP}_{-wh}))\\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\\ \lambda\rho\lambda\varphi.\rho(\lambda\varphi_0.\varphi_0)(\varphi);\\ \lambda\mathscr{F}.\mathscr{F}(\lambda u.\mathbf{talked}(u));\\ (\mathrm{VP}{\upharpoonright}\mathrm{NP}_{-wh}){\upharpoonright}((\mathrm{VP}{\upharpoonright}\mathrm{NP}_{-wh}){\upharpoonright}(\mathrm{VP}{\upharpoonright}\mathrm{NP}_{-wh}))\end{array}
}{
\lambda\varphi.\mathsf{did} \circ \varphi;\ \lambda x\lambda y.\mathbf{talk}(x)(y);\ \mathrm{VP}{\upharpoonright}\mathrm{NP}_{-wh}
}{\upharpoonright}\mathrm{E}
\quad
\left[\begin{array}{l}\varphi_3;\\ v;\\ \mathrm{NP}_{-wh}\end{array}\right]^3
}{
\mathsf{did} \circ \varphi_3;\ \lambda y.\mathbf{talked}(v)(y);\ \mathrm{VP}
}{\upharpoonright}\mathrm{E}
\quad
\begin{array}{l}\mathsf{mary};\\ \mathbf{m};\\ \mathrm{NP}\end{array}
}{
\dfrac{
\mathsf{mary} \circ \mathsf{did} \circ \varphi_3;\ \mathbf{talk}(v)(\mathbf{m});\ \mathrm{S}
}{
\lambda\varphi_3.\mathsf{mary} \circ \mathsf{did} \circ \varphi_3;\ \lambda v.\mathbf{talk}(v)(\mathbf{m});\ \mathrm{S}{\upharpoonright}\mathrm{NP}_{-wh}
}{\upharpoonright}\mathrm{I}^3
}\backslash\mathrm{E}
$$

The ellipsis operator has syntactic type schematically of the form $X{\upharpoonright}(X{\upharpoonright}X)$, and the anaphora resolution condition requires that X matches the category of the relevant antecedent whose meaning is recovered in ellipsis resolution (see Kubota and Levine (2017b) for details). In the case at hand, the appropriate antecedent is the greyed-in expression in (12), with semantics $\lambda u.\mathbf{talked}(u)$. The free variable P in the ellipsis operator thus gets resolved as this term at the step (which strictly speaking is outside of the syntactic derivation) marked as ①. But then, any attempt to compose the sign derived in (13) with the extraction operator will fail:

(14)

$$
\dfrac{
\begin{array}{c}\vdots \quad \vdots\\ \lambda\varphi_3.\mathsf{mary} \circ \mathsf{did} \circ \varphi_3;\ \lambda v.\mathbf{talk}(v)(\mathbf{m});\ \mathrm{S}{\upharpoonright}\mathrm{NP}_{-wh}\end{array}
\quad
\begin{array}{l}\lambda\sigma.\mathsf{whom} \circ \sigma(\boldsymbol{\epsilon});\\ \mathbf{wh}(\mathbf{person});\ \mathrm{Q}{\upharpoonright}(\mathrm{S}{\upharpoonright}\mathrm{NP}_{+wh})\end{array}
}{
\textbf{FAIL}
}
$$

In a nutshell, the extraction operator can only compose with a sentence missing an NP_{+wh}, but the conditions imposed on prepositions in English* allow *to* to combine only with NP_{-wh}, leading to a continuation typed $\mathrm{S}{\upharpoonright}\mathrm{NP}_{-wh}$, an invalid argument for the extraction operator. Such extractions therefore cannot give rise to well-formed VP ellipsis strandings. No special mechanisms are required, and nothing has to be stipulated other than the lexical condition, illustrated in (10), which simply expresses the ban on preposition stranding in such languages. In particular, no covert syntactic structure is required, as long as the required syntactic information about category type is made available to the anaphoric process.

In English, in contrast, sentences such as (11) are licensable because in place of (10), the lexical entry for *to* is (15):

(15) $\mathsf{to};\ \lambda x.x;\ \mathrm{PP}_{to}/\mathrm{NP}$

Since *to* can combine with an NP of either polarity for the *wh* feature, we will have a derivation along the following lines:

(16)
$$\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\begin{bmatrix}\varphi_0;\\ u;\mathrm{NP}\end{bmatrix}^1 \quad \begin{matrix}\mathsf{to};\\ \lambda\varphi.\varphi;\ \mathrm{PP}_{to}/\mathrm{NP}\end{matrix}}{\mathsf{to}\circ\varphi_0;\ u;\ \mathrm{PP}_{to}} \quad \begin{matrix}\mathsf{talked};\\ \mathbf{talk};\\ \mathrm{VP}/\mathrm{PP}_{to}\end{matrix}}{\mathsf{talked}\circ\mathsf{to}\circ\varphi_0;\ \mathbf{talk}(u);\ \mathrm{VP}}}{\lambda\varphi_0.\mathsf{talked}\circ\mathsf{to}\circ\varphi_0;\ \lambda u.\mathbf{talk}(u);\ \mathrm{VP}{\restriction}\mathrm{NP}}{\restriction}\mathrm{I}^1 \quad \begin{bmatrix}\varphi_1;\\ w;\mathrm{NP}\end{bmatrix}^2}{\mathsf{talked}\circ\mathsf{to}\circ\varphi_1;\ \mathbf{talk}(w);\ \mathrm{VP}} \quad \begin{matrix}\mathsf{john};\\ \mathbf{j};\mathrm{NP}\end{matrix}}{\mathsf{john}\circ\mathsf{talked}\circ\mathsf{to}\circ\varphi_1;\ \mathbf{talk}(w)(\mathbf{j});\ \mathrm{S}}}{\lambda\varphi_0.\mathsf{john}\circ\mathsf{talked}\circ\mathsf{to}\circ\varphi_0;\ \lambda w.\mathbf{talk}(w)(\mathbf{j});\ \mathrm{S}{\restriction}\mathrm{NP}}{\restriction}\mathrm{I}^2}{\lambda\varphi_0.\mathsf{john}\circ\mathsf{talked}\circ\mathsf{to}\circ\varphi_0;\ \lambda w.\mathbf{talk}(w)(\mathbf{j});\ \mathrm{S}{\restriction}\mathrm{NP}_{-wh}}\vee\mathrm{Lemma} \quad \begin{matrix}\mathsf{bill};\\ \mathbf{b};\mathrm{NP}_{-wh}\end{matrix}}{\mathsf{john}\circ\mathsf{talked}\circ\mathsf{to}\circ\mathsf{bill};\ \mathbf{talk}(\mathbf{b})(\mathbf{j});\ \mathrm{S}}{\restriction}\mathrm{E}$$

The greyed-in line is the critical proof step for anaphoric retrieval. The argument of this proof term belongs to the supertype NP. So, we can derive an auxiliary in the entry VP↾NP by anaphorically retrieving its meaning from this antecedent. The $+wh$ feature specification for the extracted argument comes only at a later stage, via the usual ∨Lemma in (8).

(17)
$$\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\begin{matrix}\vdots\quad\vdots\\ \lambda\sigma\lambda\varphi.\mathsf{did}\circ\sigma(\varphi);\\ \lambda f\lambda x\lambda y.f(x)(y);\\ (\mathrm{VP}{\restriction}\mathrm{NP}){\restriction}(\mathrm{VP}{\restriction}\mathrm{NP})\end{matrix} \quad \textcircled{1}\!\to \begin{matrix}\lambda\rho\lambda\varphi.\rho(\lambda\varphi_0.\varphi_0)(\varphi);\\ \lambda\mathscr{F}.\mathscr{F}(P);\\ (\mathrm{VP}{\restriction}\mathrm{NP}){\restriction}((\mathrm{VP}{\restriction}\mathrm{NP}){\restriction}(\mathrm{VP}{\restriction}\mathrm{NP}))\\ \cdots\cdots\cdots\cdots\\ \lambda\rho\lambda\varphi.\rho(\lambda\varphi_0.\varphi_0)(\varphi);\\ \lambda\mathscr{F}.\mathscr{F}(\lambda u.\mathbf{talked}(u));\\ (\mathrm{VP}{\restriction}\mathrm{NP}){\restriction}((\mathrm{VP}{\restriction}\mathrm{NP}){\restriction}(\mathrm{VP}{\restriction}\mathrm{NP}))\end{matrix}}{\lambda\varphi.\mathsf{did}\circ\varphi;\ \lambda x\lambda y.\mathbf{talk}(x)(y);\ \mathrm{VP}{\restriction}\mathrm{NP}}{\restriction}\mathrm{E} \quad \begin{bmatrix}\varphi_2;\\ w;\mathrm{NP}\end{bmatrix}^2}{\mathsf{did}\circ\varphi_2;\ \mathbf{talk}(w);\ \mathrm{VP}}{\restriction}\mathrm{E} \quad \begin{matrix}\mathsf{mary};\\ \mathbf{m};\mathrm{NP}\end{matrix}}{\mathsf{mary}\circ\mathsf{did}\circ\varphi_2;\ \mathbf{talk}(w)(\mathbf{m});\ \mathrm{S}}\backslash\mathrm{E}}{\lambda\varphi_2.\mathsf{mary}\circ\mathsf{did}\circ\varphi_2;\ \mathbf{l}w.\mathbf{talk}(w)(\mathbf{m});\ \mathrm{S}{\restriction}\mathrm{NP}}{\restriction}\mathrm{I}^2}{\lambda\varphi_2.\mathsf{mary}\circ\mathsf{did}\circ\varphi_2;\ \lambda w.\mathbf{talk}(w)(\mathbf{m});\ \mathrm{S}{\restriction}\mathrm{NP}_{+wh}}\vee\mathrm{Lemma}$$

We wish to emphasize that the difference between the failed derivation in English* and the unproblematic derivation in English is a direct reflection of the different valence possibilities for prepositions in the two language, with no reference to any structure ever coming into the picture. As the proof for English* makes clear, the enforced selection of NP_{-wh} by prepositions translates, through the chain of hypothetical reasoning displayed above, into a sign typed $\mathrm{S}{\restriction}\mathrm{NP}_{-wh}$ which is unable to compose with the extraction operator to complete the derivation. Since the $\pm wh$ distinction is an independently motivated distinction that needs to be encoded in the syntactic type of NPs, the valence information is all that is needed to institute the P-stranding generalization for the 'extraction out of VP ellipsis' pattern. We now demonstrate that exactly the same purely valence-based account is sufficient for the P-stranding generalization in the case of sluicing as well.

Sluicing Following the overall strategy in Barker (2013), we posit the following operator for the analysis of sluicing:

(18) $\lambda\rho.\rho(\lambda\varphi.\varphi);\ \lambda\mathscr{W}.\mathscr{W}(P);\ \mathrm{Q}{\restriction}(\mathrm{Q}{\restriction}(\mathrm{S}{\restriction}\mathrm{NP}_{+wh}))$
——where P is a property matching a contextually salient sign compatible with the type description $\mathrm{S}{\restriction}\mathrm{NP}_{+wh}$ in the preceding discourse

In simple cases of sluicing, we obtain derivations such as that given in (20) for (19):

(19) John criticized someone, but Mary doesn't know whom.

(20)

$$\dfrac{\text{criticized}; \mathbf{criticize}; \mathrm{VP/NP} \quad [\varphi_0; x; \mathrm{NP}]^1}{\text{criticized} \circ \varphi_0; \mathbf{criticize}(x); \mathrm{VP}} \quad \text{john}; \mathbf{j}; \mathrm{NP}$$

$$\dfrac{\text{criticized} \circ \varphi_0; \mathbf{criticize}(x); \mathrm{VP} \quad \text{john}; \mathbf{j}; \mathrm{NP}}{\text{john} \circ \text{criticized} \circ \varphi_0; \mathbf{criticize}(x)(\mathbf{j}); \mathrm{S}}$$

$$\dfrac{\text{john} \circ \text{criticized} \circ \varphi_0; \mathbf{criticize}(x)(\mathbf{j}); \mathrm{S}}{\lambda\varphi_0.\text{john} \circ \text{criticized} \circ \varphi_0; \lambda x.\mathbf{criticize}(x)(\mathbf{j}); \mathrm{S}{\upharpoonright}\mathrm{NP}}{\upharpoonright}\mathrm{I}^1$$

$$\dfrac{\lambda\varphi_0.\text{john} \circ \text{criticized} \circ \varphi_0; \lambda x.\mathbf{criticize}(x)(\mathbf{j}); \mathrm{S}{\upharpoonright}\mathrm{NP}}{\lambda\varphi_0.\text{john} \circ \text{criticized} \circ \varphi_0; \lambda x.\mathbf{criticize}(x)(\mathbf{j}); \mathrm{S}{\upharpoonright}\mathrm{NP}_{-wh}} \vee\mathrm{Lemma}$$

$$\dfrac{\lambda\varphi_0.\text{john} \circ \text{criticized} \circ \varphi_0; \lambda x.\mathbf{criticize}(x)(\mathbf{j}); \mathrm{S}{\upharpoonright}\mathrm{NP}_{-wh} \quad \lambda\sigma_0.\sigma_0(\text{someone}); \mathbf{\exists_{person}}; \mathrm{S}{\upharpoonright}(\mathrm{S}{\upharpoonright}\mathrm{NP}_{-wh})}{\text{john} \circ \text{criticized} \circ \text{someone}; \mathbf{\exists_{person}}(\lambda x.\mathbf{criticize}(x)(\mathbf{j})); \mathrm{S}}$$

$$\dfrac{\lambda\rho.\rho(\lambda\varphi.\varphi); \lambda\mathscr{W}.\mathscr{W}(P); \mathrm{Q}{\upharpoonright}(\mathrm{Q}{\upharpoonright}(\mathrm{S}{\upharpoonright}\mathrm{NP}_{+wh}))}{\lambda\rho.\rho(\lambda\varphi.\varphi); \lambda\mathscr{W}.\mathscr{W}(\lambda x.\mathbf{criticize}(x)(\mathbf{j})); \mathrm{Q}{\upharpoonright}(\mathrm{Q}{\upharpoonright}(\mathrm{S}{\upharpoonright}\mathrm{NP}_{+wh}))}$$

$$\dfrac{\lambda\rho.\rho(\lambda\varphi.\varphi); \lambda\mathscr{W}.\mathscr{W}(\lambda x.\mathbf{criticize}(x)(\mathbf{j})); \mathrm{Q}{\upharpoonright}(\mathrm{Q}{\upharpoonright}(\mathrm{S}{\upharpoonright}\mathrm{NP}_{+wh})) \quad \lambda\sigma_1.\text{whom} \circ \sigma_1(\epsilon); \mathbf{wh}(\mathbf{person}); \mathrm{Q}{\upharpoonright}(\mathrm{S}{\upharpoonright}\mathrm{NP}_{+wh})}{\text{whom}; \mathbf{wh}(\mathbf{person})(\lambda x.\mathbf{criticize}(x)(\mathbf{j})); \mathrm{Q}}$$

$$\dfrac{\text{whom}; \mathbf{wh}(\mathbf{person})(\lambda x.\mathbf{criticize}(x)(\mathbf{j})); \mathrm{Q} \quad \text{know}; \mathbf{know}; \mathrm{VP/Q}}{\text{know} \circ \text{whom}; \mathbf{know}(\mathbf{wh}(\mathbf{person})(\lambda x.\mathbf{criticize}(x)(\mathbf{j}))); \mathrm{VP}}$$

$$\dfrac{\text{know} \circ \text{whom}; \mathbf{know}(\mathbf{wh}(\mathbf{person})(\lambda x.\mathbf{criticize}(x)(\mathbf{j}))); \mathrm{VP} \quad \text{doesn't}; \lambda Q\lambda y.\neg Q.Q(y); \mathrm{VP/VP}}{\text{doesn't} \circ \text{know} \circ \text{whom}; \lambda y.\neg\mathbf{know}(\mathbf{wh}(\mathbf{person})(\lambda x.\mathbf{criticize}(x)(\mathbf{j})))(y); \mathrm{VP}}$$

$$\dfrac{\text{doesn't} \circ \text{know} \circ \text{whom}; \lambda y.\neg\mathbf{know}(\mathbf{wh}(\mathbf{person})(\lambda x.\mathbf{criticize}(x)(\mathbf{j})))(y); \mathrm{VP} \quad \text{mary}; \mathbf{m}; \mathrm{NP}}{\text{mary} \circ \text{doesn't} \circ \text{know} \circ \text{whom}; \neg\mathbf{know}(\mathbf{wh}(\mathbf{person})(\lambda x.\mathbf{criticize}(x)(\mathbf{j})))(\mathbf{m}); \mathrm{S}}$$

The lexical entry for the *wh*-word *whom* and the sluicing operator (18) are taken to be common to English* and English, and more generally, to all languages with *wh*-extraction and sluicing regardless of whether they allow stranded prepositions. Again, the sole difference is in the specification of the class of NPs with which prepositions can combine.

Given the entry for *to* in (10), preposition stranding is already automatically blocked in English* sluicing, just as it is for 'VPE extraction'. For example, consider the following example:

(21) John talked to someone, but I don't know who(m).

In order to create the appropriate expression to serve as the antecedent in the first clause, *to* needs to combines with a variable, but in view of (10), this variable will necessarily be $-wh$. When the hypothesis corresponding to this variable is withdrawn, the result, $\mathrm{S}{\upharpoonright}\mathrm{NP}_{-wh}$, will be unable to serve as an antecedent for the sluicing operator, which explicitly requires the antecedent to be of type $\mathrm{S}{\upharpoonright}\mathrm{NP}_{+wh}$.

(22)

$$\dfrac{\text{to}; \lambda\varphi.\varphi; \mathrm{PP}_{to}/\mathrm{NP}_{-wh} \quad [\varphi_1; v; \mathrm{NP}_{-wh}]^1}{\text{to} \circ \varphi_1; v; \mathrm{PP}}$$

$$\dfrac{\text{talked}; \mathbf{talk}; \mathrm{VP/PP}_{to} \quad \text{to} \circ \varphi_1; v; \mathrm{PP}}{\text{talked} \circ \text{to} \circ \varphi_1; \mathbf{talk}(v); \mathrm{VP}}$$

$$\dfrac{\text{talked} \circ \text{to} \circ \varphi_1; \mathbf{talk}(v); \mathrm{VP} \quad \text{john}; \mathbf{j}; \mathrm{NP}}{\text{john} \circ \text{talked} \circ \text{to} \circ \varphi_1; \mathbf{talk}(v)(\mathbf{j}); \mathrm{S}}$$

$$\dfrac{\text{john} \circ \text{talked} \circ \text{to} \circ \varphi_1; \mathbf{talk}(v)(\mathbf{j}); \mathrm{S}}{\lambda\varphi_1.\text{john} \circ \text{talked} \circ \text{to} \circ \varphi_1; \lambda v.\mathbf{talk}(v)(\mathbf{j}); \mathrm{S}{\upharpoonright}\mathrm{NP}_{-wh}}{\upharpoonright}\mathrm{I}^1$$

$$\dfrac{\lambda\sigma_1.\sigma_1(\text{someone}); \mathbf{\exists_{person}}; \mathrm{S}{\upharpoonright}(\mathrm{S}{\upharpoonright}\mathrm{NP}_{-wh}) \quad \lambda\varphi_1.\text{john} \circ \text{talked} \circ \text{to} \circ \varphi_1; \lambda v.\mathbf{talk}(v)(\mathbf{j}); \mathrm{S}{\upharpoonright}\mathrm{NP}_{-wh}}{\text{john} \circ \text{talked} \circ \text{to} \circ \text{someone}; \mathbf{\exists_{person}}(\lambda v.\mathbf{talk}(v)(\mathbf{j})); \mathrm{S}}$$

$$\lambda\rho.\rho(\lambda\varphi.\varphi);\ \lambda\mathscr{W}.\mathscr{W}(P);\ \mathrm{Q}{\upharpoonright}(\mathrm{Q}{\upharpoonright}(\mathrm{S}{\upharpoonright}\mathrm{NP}_{+wh}))$$

$$\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots$$

$$\lambda\rho.\rho(\lambda\varphi.\varphi);\ \lambda\mathscr{W}.\mathscr{W}(\lambda v.\mathbf{talk}(v)(\mathbf{j}));\ \mathrm{Q}{\upharpoonright}(\mathrm{Q}{\upharpoonright}(\mathrm{S}{\upharpoonright}\mathrm{NP}_{+wh}))\quad \textbf{FAIL}\quad \lambda\sigma.\mathsf{who}\circ\sigma(\epsilon);\ \mathbf{wh}(\mathbf{person});\mathrm{Q}{\upharpoonright}(\mathrm{S}{\upharpoonright}\mathrm{NP}_{+wh})$$

$$\overline{\mathsf{who};\ \mathbf{wh}(\mathbf{person})(\lambda v.\mathbf{talk}(v)(\mathbf{j}));\ \mathrm{Q}}$$

By contrast, in English, since the argument of *to* is underspecified for the *wh* feature, the following derivation is available:

(23)

$$\mathsf{to};\ \lambda\varphi.\varphi;\ \mathrm{PP}_{to}/\mathrm{NP}\quad \varphi_1;\ v;\ \mathrm{NP}$$

$$\mathsf{talked};\ \mathbf{talk};\ \mathrm{VP}/\mathrm{PP}_{to}\quad \overline{\mathsf{to}\circ\varphi_1;\ v;\ \mathrm{PP}}$$

$$\overline{\mathsf{talked}\circ\mathsf{to}\circ\varphi_1;\ \mathbf{talk}(v);\ \mathrm{VP}}\quad \mathsf{john};\ \mathbf{j};\ \mathrm{NP}$$

$$\overline{\mathsf{john}\circ\mathsf{talked}\circ\mathsf{to}\circ\varphi_1;\ \mathbf{talk}(v)(\mathbf{j});\ \mathrm{S}}$$

$$\overline{\lambda\varphi_1.\mathsf{john}\circ\mathsf{talked}\circ\mathsf{to}\circ\varphi_1;\ \lambda v.\mathbf{talk}(v)(\mathbf{j});\ \mathrm{S}{\upharpoonright}\mathrm{NP}}$$

$$\lambda\sigma_1.\sigma_1(\mathsf{someone});\ \mathbf{\exists_{person}};\mathrm{S}{\upharpoonright}(\mathrm{S}{\upharpoonright}\mathrm{NP}_{-wh})\qquad \overline{\lambda\varphi_1.\mathsf{john}\circ\mathsf{talked}\circ\mathsf{to}\circ\varphi_1;\ \lambda v.\mathbf{talk}(v)(\mathbf{j});\ \mathrm{S}{\upharpoonright}\mathrm{NP}_{-wh}}\ \vee\mathrm{Lemma}$$

$$\overline{\mathsf{john}\circ\mathsf{talked}\circ\mathsf{to}\circ\mathsf{someone};\ \mathbf{\exists_{person}}(\lambda v.\mathbf{talk}(v)(\mathbf{j}));\ \mathrm{S}}$$

$$\lambda\rho.\rho(\lambda\varphi.\varphi);\ \mathbf{wh}(\mathbf{person})(P);\mathrm{Q}{\upharpoonright}(\mathrm{Q}{\upharpoonright}(\mathrm{S}{\upharpoonright}\mathrm{NP}_{+wh}))$$

$$\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots$$

$$\lambda\rho.\rho(\lambda\varphi.\varphi);\ \mathbf{wh}(\mathbf{person})(\lambda v.\mathbf{talk}(v)(\mathbf{j}));\mathrm{Q}{\upharpoonright}(\mathrm{Q}{\upharpoonright}(\mathrm{S}{\upharpoonright}\mathrm{NP}_{+wh}))\quad \lambda\sigma.\mathsf{whom}\circ\sigma(\epsilon);\ \mathbf{wh}(\mathbf{person});\mathrm{Q}{\upharpoonright}(\mathrm{S}{\upharpoonright}\mathrm{NP}_{+wh})$$

$$\overline{\mathsf{whom};\ \mathbf{wh}(\mathbf{person})(\lambda v.\mathbf{talk}(v)(\mathbf{j}));\ \mathrm{Q}}$$

Here, the free variable P in the sluicing operator can be instantiated as a contextually appropriate predicate $\lambda v.\mathbf{talk}(v)(\mathbf{j})$ denoted by the greyed-in expression of type $\mathrm{S}{\upharpoonright}\mathrm{NP}$ in the antecedent clause, since $\mathrm{S}{\upharpoonright}\mathrm{NP}$, which entails $\mathrm{S}{\upharpoonright}\mathrm{NP}_{+wh}$ via the $\vee$Lemma in (8) (i.e. $\mathrm{S}{\upharpoonright}\mathrm{NP} \vdash \mathrm{S}{\upharpoonright}\mathrm{NP}_{+wh}$), is clearly compatible with the description $\mathrm{S}{\upharpoonright}\mathrm{NP}_{+wh}$. It then follows that in English, preposition stranding is possible in sluicing.

Hence the P-stranding generalization in the case of sluicing falls out directly from the simple lexical treatment in (10) and (18), with no need to posit covert configurations.

Conclusion The technical details of the proofs given above are necessary so that readers can verify for themselves that our proposal does exactly what we claim it does: guarantee that the P-stranding generalization does indeed fall out of our type-logical framework with no appeal whatever to configurational representations characterizing the 'missing' material in any ellipsis construction. We hope, however, that the fundamental simplicity of our solution will not become lost in these technical details. The central point is that in order to capture the P-stranding generalization, nothing more need be assumed than the independently needed lexical prohibition on NP_{+wh} arguments to prepositions in non-stranding languages and the independently motivated assumption that the anaphora recovery process in ellipsis is sensitive to the syntactic category of the antecedent expression (cf. Kubota and Levine 2017a,b).

Importantly, access to syntactic category information that our anaphora-based approach crucially exploits is not something that is 'added on' to the theory, but is a direct consequence of the fundamental architecture of (most versions of) categorial grammar: at each stage of syntactic derivation, the prosody, semantics and the syntactic type of the linguistic expression are fully explicit. The correlation between a particular semantics and specific syntactic type is thus built into the fundamental architecture of the theory. This architecture however does not allow access to the 'history of derivation' (i.e., the structure of the proof) up to the point that the expression in question is obtained, and it is in this respect that categorial grammar departs most crucially from derivational variants of syntax that in principle allow (unless additional theory-internal assumptions are made) full access to the internal syntactic structure of a linguistic expression.

Both H-TLCG and the P&P analyses of ellipsis essentially rely on specifications of lexical valence to rule out overgeneration that would arise in purely interpretive accounts, and in this respect, at the descriptive level both are getting at more or less the same insight. However, covert structure analyses add a further component of hierarchical representations projected from these lexically specified argument structure possibilities—representations which, given the foregoing discussion, are not necessary to capture the P-stranding generalization. Basic considerations of parsimony (i.e. Occam's razor) thus seem to rule in favor of the H-TLCG account (unless it can be shown that this approach incurs some hidden or overlooked additional complexity that is not present in the derivational approach), and the view that syntactic information as reflected in the syntactic categories of linguistic expressions is sufficient in the licensing of elliptical constructions, without the need for hidden phrase structure.

References

Barker, Chris. 2013. Scopability and sluicing. *Linguistics and Philosophy* 36(3):187–223.

Bayer, Samuel. 1996. The coordination of unlike categories. *Language* 72:579–616.

Bayer, Samuel and Mark Johnson. 1995. Features and agreement. In *Proceedings of ACL 33*, 70–76. Boston: Association for Computational Linguistics.

Karttunen, Lauri. 1977. Syntax and semantics of questions. *Linguistics and Philosophy* 1(1):3–44.

Kubota, Yusuke and Robert Levine. 2014. Hybrid Type-Logical Categorial Grammar. ms., University of Tsukuba and Ohio State University, available at `http://ling.auf.net/lingbuzz/002313`.

Kubota, Yusuke and Robert Levine. 2017a. Ellipsis and syntactic information. University of Tsukuba/Ohio State University ms.; available at http://ling.auf.net/lingbuzz/003643.

Kubota, Yusuke and Robert Levine. 2017b. Pseudogapping as pseudo-VP ellipsis. *Linguistic Inquiry* 48(2):213–257.

Merchant, Jason. 2001. *The Syntax of Silence*. Oxford: OUP.

Merchant, Jason. 2013. Ellipsis: A survey of analytical approaches. ms., University of Chicago.

Morrill, Glyn. 1994. *Type Logical Grammar: Categorial Logic of Signs*. Dordrecht: Kluwer.

Muskens, Reinhard. 2003. Language, lambdas, and logic. In G.-J. Kruijff and R. Oehrle, eds., *Resource Sensitivity in Binding and Anaphora*, 23–54. Dordrecht: Kluwer.

Nykiel, Joanna. 2013. Clefts and preposition omission under sluicing. *Lingua* 123:74–117.

Sag, Ivan and Joanna Nykiel. 2011. Remarks on sluicing. In S. Müller, ed., *The Proceedings of the 18th International Conference on Head-Driven Phrase Structure Grammar*, 188–208. Stanford: CSLI Publications.

Steedman, Mark. 1996. *Surface Structure and Interpretation*. Cambridge, Mass.: MIT Press.

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