On Internal Merge

Mark Steedman September 7, 2023

Abstract The rule MOVE, used in various forms in generative grammars to capture displacement or discontinuous constituency, has recently been talked of as an "internal" version of MERGE, the operation of simple node- or set-formation. Internal merge "reconstructs" the displaced element in its original argument-structural position at the level of logical form via a "copy", to which it has been identical throughout the derivation.

Reducing MOVE to MERGE seems to be on the side of simplifying the theory of grammar, potentially eliminating the need for constraints on movement in order to limit overgeneration. The paper addresses the question of how internal merge should be defined in formal terms.

An account of discontinuity is proposed in which copies originate in the lexicon, as seems to be required by a strict interpretations of the Inclusiveness Condition of Chomsky (1995b), where they can be thought of as binders and variables in lexical logical form (lf). Merger is defined via a small number of type-dependent combinatory rules, which apply to strictly string-adjacent categories to monotonically project from the lexical array varieties of discontinuous dependencies that have been described in terms of various forms of movement, including "A", "$\overline{\text{A}}$", "remnant", "head", "parallel", "sideward", "covert", "roll-up", and "late merge", without any attendant "constraints on movement" other than those projected from lexical types. The analysis extends to a plethora of other discontinuous operations that have been proposed in addition to or instead of MOVE, including AGREE, LABEL, TRANSFER, and DELETE, all of which are replaced by synchronous monotonic lf and pf merger of contiguous categories. The result is to eliminate structure-dependence and action-at-a-distance of all kinds from syntactic rules.[*]

1 Introduction

The Inclusiveness Condition of Chomsky (1995b:228, 2001:2, 2001/2004:109) prohibits rules of syntactic derivation from adding any information such as "indices, traces, syntactic categories or bar-levels, and so on" to that which has already been specified in the numeration or multiset of categories with which the derivation begins.

This condition, and the related "No-Tampering" and "Extension" Conditions requiring monotonicity in rules, are sometimes sidelined as not-obviously-substantive constraints on notation, or as "[desiderata] of efficient processing" (Chomsky, 2008:138, 1993; 1995b:189-91; Adger, 2003:95-96). In the present paper, they are taken to constitute core characteristics of a radically lexicalized theory of grammar, requiring that all language-specific details of combinatory potential, such as category, subcategorization, agreement, linearization, and the like, be specified at the level of the lexicon, and thereafter be only either "checked" or "projected" unchanged onto derived categories by universal, language-independent, type-dependent rules.

The prevalence in all natural languages of discontinuous dependencies like the following,

---

[*] Earlier versions of this paper circulated under the title "Projecting Dependency". Thanks to David Adger, Cem Bozşahin, Caroline Heycock, Julia Hockenmaier, Tim Hunter, Polly Jacobson, Bob Levine, Haixia Man, Elise Newman, Geoff Pullum, Miloš Stanojević, John Torr, Rob Truswell, Bonnie Webber, and the considerate reviewers for LI. The project SEMANTAX has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement No. 742137).

where elements that belong together in the meaning representation are separated from each other in the sentence, seems to present a major challenge to this reasonable assumption:

(1) a. Mary$_i$ was fired$_i$.
b. Mary$_i$ seems to sleep$_i$.
c. Mary$_i$ I think I like$_i$.

Discontinuities of this kind were originally analysed by Chomsky in terms of rules of movement, relating elements across potentially unbounded structures. Modern Minimalist versions of the theory distinguish two fundamental derivational operations: MERGE (sometimes distinguished as "external" merge), the simple combination of constituents to form a set; and MOVE (sometimes distinguished as "internal" merge), relating discontiguous elements realized as "copies". Copies are held to be identical throughout the derivation (Epstein et al., 1998; Chomsky, 2001/2004, 2019; Adger, 2003; Epstein and Seely, 2006), such that merger with one simultaneously instantiates the other.

More recently, this analysis has been further complicated by the introduction of (at least) the following asynchronous operations in addition to, or as replacements for, external/internal MERGE (Chomsky, 2000, 2008): long-distance AGREE (sometimes as a mechanism for internal merge); LABEL (in place of X-bar theory); TRANSFER of intermediate representations to PF and LF "interfaces"; DELETE (possibly as a component of transfer); and a criterion for distinguishing copies arising from movement from genuine repetition, possibly via the NUMERATION, or multiset of lexical items as input to syntactic derivation (Collins and Groat, 2018).

This paper inquires how internal merger can be formally unified with the standard external variety without making discontinuity an operation of the theory itself. Displacement is taken to be the phenomenon to be explained, using as little expressive power as possible beyond that of the context-free grammars that seem to be all that is needed to cover first-order logical forms and programming language syntax.

There have been other proposals for eliminating movement, such as: Generalized/Head-driven Phrase Structure Grammar (GPSG/HPSG)-style trace-feature passing (Gazdar, 1981; Gazdar et al., 1985; Neeleman and van de Koot, 2010); Categorial Grammar (CG)-style nonstandard constituency (Ades and Steedman, 1982; Williams, 2003); Lexical-Functional Grammar (LFG)-style lexicalization of locality (Bresnan, 1982; Borer, 1984; Manzini, 1992; Brody, 1995); HPSG/LFG-style structural unification (Kay, 1984; Pollard and Sag, 1994; Bresnan, 2001); various base-generative anaphoric proposals (Koster, 1978; Adger and Ramchand, 2005; Jacobson, 2014); or mediation via various Transition Network Grammar (TNG)-style HOLD registers, MOVE-boxes, or the numeration itself (Woods, 1970; Fong, 2005; Stroik, 2009).

However, in practice such mechanisms have tended either to be so restrictive as to cause the grammar to collapse to the descriptively-inadequate context-free class (CFG, as GPSG was avowedly, and as Rogers (1998:185) showed to be the case for a version of Government-Binding theory limited by Rizzi's 1990 Relativized Minimality and Manzini's lexicon, and as Kobele, 2010 shows for a related version of Minimalism), or to make the theory overly expressive (Peters and Ritchie, 1973; Gärtner and Michaelis, 2007). These alternatives typically required stipulation of ad hoc substantive constraints to limit over-generalization, such as "shortest move" or "path containment" conditions (Rosenbaum, 1967; Pesetsky, 1982; Kayne, 1983; Rizzi, 1990; Chomsky, 1995b), many of which had to be replicated in the lexicon (Marantz, 1997). Attempts in the other direction to augment context-free theories such as GPSG or TNG to cover multiple discontinuities showed a similar tendency to explode expressive power and compromise explanatory adequacy.

If the inclusiveness and no-tampering conditions are core principles, then the relation between the pairs of structural positions corresponding to the indices in (1a-c) must be established before the derivation begins, in the lexicon or the numeration derived from it. Under the present proposal, this relation is expressed at the level of lexical logical form as a binder $\lambda\alpha$ and its variable α. As the copy theory of movement requires (Epstein and Seely, 2006:16,n.2,32; Chomsky, 2007:10), the two αs are literally identical: whatever value $\lambda\alpha$ combines with is the value of every occurrence of α. In that sense, "Reconstruction" of α at the source(s) of movement follows immediately (von Stechow, 1991:133). In the same sense, internal merge is not distinct from external merge, since external merge with a lexicalized element *Mary* in (1a-c) will identically instantiate the copy with which it has been lexically λ-entangled from the start.

The present paper will show that the effects of movement/internal merge can be derived from a lexicalized theory of grammar of low "near-context-free" expressive power, in which categories are defined syntactically and semantically as functions and/or arguments, and all discontinuities are base-generated in the logical forms associated with lexical items such as raising verbs and *wh*-words, using the same mechanism that earlier theories have used to treat local selection.

The syntactic operations that then project lexical dependencies onto sentences are limited to strictly adjacent merger according to pure functional operations such as functional application and functional composition, without attendant constraints other than those originating in lexical specification. Such a theory constitutes an "optimal derivational system" (Epstein and Seely:178-179,n.6), in the sense that it is formally impossible to condition such rules on derivations or structural representations, including logical form (lf, which in this paper is distinguished from other notions of Logical Form (LF) by the use of lower case).

The remainder of the paper examines a sequence of core examples of major discontinuous constructions that have proved challenging to the standard theories of movement, numerous constraints upon which are shown to be redundant under the present approach. Otherwise, the aims of the paper are severely limited. Nothing is said about such constructions as Extraposition, Sluicing, VP-Ellipsis, and SVO Gapping, which the paper follows Emonds (1979), Ginzburg and Sag (2000), and Hardt and Romero (2004), among others, in regarding as mediated at least in part by discourse anaphora and/or co-reference. For the same reason, while section 6.2 touches on the consequences of copy-based reconstruction for binding theory, rather little will be said otherwise about the mechanism by which pronouns actually obtain their values.

The paper begins in section 2 by defining a categorial "Bare Phrase Structural" notation for linguistic categories and the universal rules of contiguous merger by which they combine, including function composition, which are shown to be non- context-free. Section 3 then analyses Raising and *There*-insertion in terms of application and function composition merger alone, without the need for derivational constraints such as the Θ-criterion. Section 4 concerns "scrambling" and word-order variation in Germanic, where the effects of Wallenberg's 2009 constraints are similarly predicted, and introduces the categorial realization of Case as morpho-lexical "type-raising", which gives arguments the lexical category of functions over the lexical frames that subcategorize for them, thereby defining the domain of scrambling. Section 5 analyses Coordination, showing the Coordinate Structure Constraint and its ramifications to be predicted. Section 6 analyses *Wh*-extraction, addressing certain "Late Merge" constructions, together with the prediction of Island effects for adjuncts first noted by Cattell (1976), including a brief discussion of some cases of apparently cyclic *wh*-movement in Celtic languages that might be thought to controvert the present proposal. Section 7 then considers some exceptional asymmetries between leftward and rightward extraction, including the Fixed Subject Effect and its exceptions (Bresnan, 1972), and the Right Edge Restriction (Wilder, 1999), The concluding section 8 discusses the combinatory theory more generally, including the trade-off between the elimination of discontinuity and derivation-dependence in rules, and the extended definition of derivational constituency engendered by so doing. Under this definition, all syntactic discontinuous dependency is reduced to the merger of a constituent constituting a second-order function with a contiguous constituent constituting a first-order function or property. The conclusion rejects the need for action-at-a-distance and structure-dependence in rules of grammar. Once the Inclusiveness Condition is taken seriously, lexicalization and locality of selection is all that is needed.

2 Adjacent Type-dependent Merger and the Inclusiveness Condition

The following assumptions are held in what follows to be constitutive of the theory of grammar:

(2) *The Categorial Assumption*: Linguistic Categories are defined syntactically and semantically as FUNCTIONS and/or ARGUMENTS;

(3) *The Adjacency Assumption*: Rules are pure FUNCTIONAL binary operations, limited to APPLICATION, COMPOSITION, and SUBSTITUTION, applying to strictly adjacent, phonologically-realized categories, which synchronously and monotonically compose logical forms (lf) and concatenate phonological forms (pf).

It will become apparent that the Minimalist conditions of Inclusiveness, Extension, and No-tampering follow immediately from these two assumptions, as formal universals, or theorems characterizing the Computation, as opposed to substantive constraints.[1]

2.1 Categories, Linearization, and Local Agreement

If we think of predicates (such as *works*) as functions, and their subjects and complements (such as as *Egon*) as arguments, then the simplest case of merger amounts to application of a function to an argument to yield a result (such as *Egon works*). We might provisionally represent such a derivation as follows, using the notation of Categorial Grammar (CG, Ajdukiewicz, 1935; Bar-Hillel, 1953; Lambek, 1958; Ades and Steedman, 1982; Steedman, 1987), where *works* has the type $S\backslash NP$ of an intransitive verb or function from NPs to its left into sentences S, underline indicates merger, and $<$ identifies the merger as backward application, or contiguous merger with leftward linearization:[2]

(4)
$$\begin{array}{cc} \text{Egon} & \text{works} \\ \overline{NP_{3s}} & \overline{S\backslash NP_{3s}} \\ \hline \multicolumn{2}{c}{S} \end{array}{<}$$

Subscripts *3s* and *3* informally represent bundles of agreement features for attributes including *person* and *number*, with the latter bundle underspecified for number, meaning that it is compat-

---

[1] Substitution rules (Szabolcsi, 1983/1989; Steedman, 1987, 1996) will be ignored in this paper.

[2] We adopt the "result leftmost" convention of Ajdukiewicz for categorial slashes, rather than the "result topmost" convention of Bar-Hillel and Lambek used in "Type-Logical" CG (Moortgat, 1988; Morrill, 2011).

ible with any value on that attribute. Such bundles can merge or "unify" (Shieber, 1986; Adger, 2010) just in case the values on all attributes are equal, or one is less specified than the other. In the latter case, the merged value on that attribute will be the more specified one.[3]

Similarly, we have the following derivation, in which ">": indicates forward application or rightward-linearized contiguous merge, and "<" the backward/leftward version (*agr* is an unspecified agreement feature compatible with any value):[4]

(5)

$$\begin{array}{ccccc}
\text{A} & \text{woman} & \text{saw} & \text{a} & \text{cat}\\
\hline
NP_{3s}/N_{3s} & N_{3s} & (S\backslash NP_{agr})/NP & NP_{3s}/N_{3s} & N_{3s}\\
\cline{1-2}\cline{4-5}
\multicolumn{2}{c}{NP_{3s} \quad >} & & \multicolumn{2}{c}{NP_{3s} \quad >}\\
\cline{3-5}
& & \multicolumn{3}{c}{S\backslash NP_{agr} \quad >}\\
\hline
\multicolumn{5}{c}{S \quad <}
\end{array}$$

CG lexical categories like that of *works*, $S\backslash NP_{3s}$, and *saw*, $(S\backslash NP_{agr})/NP$, are of course comparable to lexical categories in "Bare Phrase Structural" Minimalism (Chomsky, 1995a, 2001), such as the following, in which "*u*N" (for "uninterpretable N-feature") takes the place of both "$/NP$" and "$\backslash NP$" (Adger, 2003:86):

(6) work [V, *u*N] ("yields V; selects N")

(7) see [V, *u*N, *u*N] ("yields V; selects two N")

"Uninterpretable" features such as *u*N must be "checked" against or "canceled" by matching "interpretable" features such as N, carried by their arguments, a process which corresponds to matching of $/NP$ and NP under function application in derivations (4) and (5).

Minimalism can therefore be seen as a form of Categorial Grammar, with the addition of derivation-dependence and discontinuity in rules such as MOVE/INTERNAL MERGE (Berwick and Epstein, 1995a,b; Adger, 2003, 2010, 2013; Smith and Cormack, 2015), both of which are absent from the present theory.

An important further difference is that CG categories specify linearization order in language-specific lexical categories via the slash notation.[5]

Crucially for the derivation below of Inclusiveness as a theorem, the direction of specification of an argument $/Y$ or $\backslash Y$ is formally a feature-value of Y comparable to agreement 3*s* or *agr*, grounded via variables over string-position of the functor X/Y ($X\backslash Y$), requiring that its right (left) edge equals the left (right) edge of its argument Y.[6]

---

[3] It might seem more standard in Minimalist terms to write these categories as *DP* and $TP\backslash DP$, and the result of their merger as *TP*, deriving $TP\backslash DP$ from a tensed head *T* and the stem *work* := $vP\backslash DP$. However, it will become apparent that much of the motivation for the T/D notation is obviated by the form of lexicalism and Bare Phrase Structure (Chomsky, 1995b,a) that is implicit in Categorial Grammars. See also Bruening (2009, 2020); Chomsky (2019) and Pullum and Miller (2022) on the DP hypothesis.

[4] It will become apparent later that NP arguments like subjects (and therefore, determiners) are morpho-lexically type-raised, to become functions over predicates like verbs, and that the merger in (4) is actually forward combination of *Egon* with *works* to yield the same syntactic and semantic result. For the moment we can ignore this complication.

[5] The categorially-influenced Minimalist Grammars of Harkema (2001) and Torr (2019); Torr et al. (2019) also lexicalize linearity. Linearization of categories and rules is a source of strength in the present theory, for example in predicting the dependency of island effects and deletion under coordination on basic word-order (Ross, 1967, 1970).

[6] See Steedman, 1991b, 2000b:213-224 for the unification-based details, which we pass over here.
It is important to be clear that this is a quite different interpretation of the slash notation to that of GPSG/HPSG, defining directional selection, rather than extraction per se (Gazdar, 1981:159).

Derivations like (4) and (5) also build logical forms in synchronous lock-step, as in the following version of (4), in which a logical form is associated with each lexical category via the separator ":". Concatenation $f\,a$ in lf denotes application of f to a, and "associates to the left", so that $f\,a\,b$ is equivalent to $(f\,a)\,b$, while the variable-binding λ operator is used to merge values like *egon* for the NPs with the argument-structure within the logical form of the verb, via variables like y:[7]

(8)
$$\begin{array}{cc}
\text{Egon} & \text{works}\\
\overline{NP_{3s}: egon} & \overline{S\backslash NP_{3s}: \lambda y.pres(work\,y)}\\
\hline
\multicolumn{2}{c}{S: pres(work\,egon)} \;<
\end{array}$$

Similarly:[8]

(9)
$$\begin{array}{ccccc}
\text{A} & \text{woman} & \text{saw} & \text{a} & \text{cat}\\
\overline{NP_{3s}/N_{3s}: a} & \overline{N_{3s}: woman} & \overline{(S\backslash NP_{agr})/NP: \lambda x\lambda y.past(see\,x\,y)} & \overline{NP_{3s}/N_{3s}: a} & \overline{N_{3s}: cat}\\
\multicolumn{2}{c}{\overline{NP_{3s}: a\,woman}\;>} & & \multicolumn{2}{c}{\overline{NP_{3s}: a\,cat}\;>}\\
 & & \multicolumn{3}{c}{\overline{S\backslash NP_{3s}: \lambda y.past(see(a\,cat)\,y)}\;>}\\
\multicolumn{5}{c}{\overline{S: past(see(a\,cat)(a\,woman))}\;<}
\end{array}$$

In the above English examples, λ-binding happens to be redundant, since the order of command that seems to be universally required in predicate-argument structure at lf by the binding theory allows us to assume it is respected (as the lexical substitution test implies) in the order of lexical syntactic SVO selection. However, such is not the case in other languages. For example, in VSO languages such as Welsh, λ-binding wraps the first (subject) argument around the second (object) argument at the level of lexical lf, so that, as in English and by assumption every other language, the former commands the latter in the predicate-argument structure.[9]

(10)
$$\begin{array}{ccc}
\text{Gwelodd} & \text{dynes} & \text{gath}\\
\overline{(S/NP)/NP: \lambda y\lambda x.past(see\,x\,y)} & \overline{NP: a\,woman} & \overline{NP: a\,cat}\\
\multicolumn{2}{c}{\overline{S/NP: \lambda x.past(see\,x(a\,woman))}\;>} & \\
\multicolumn{3}{c}{\overline{S: past(see(a\,cat)(a\,woman))}\;>}
\end{array}$$

As far as the syntactic derivations go, apart from the inclusion of lexical linearization, and the consequent elimination of head-movement from syntactic derivation, the above fragments of English and Welsh resemble a Minimalist analysis confined to External Merger alone. Any discontinuity of vP-internal predicate-argument structural dominance relations is projected from lf via the lexical category. The external merger case of Chomsky's "labeling algorithm" (2008:145)

---

[7] Again, the verb could be written in more standard Minimalist terms as $TP\backslash DP: \lambda y.pres(work\,y)$, derived via present tense morphology from *work*, $vP: work$. Lest it be thought that we are smuggling in movement under cover of λ-abstraction, it should be noted that the use of λ-bound variables is in every case lexically defined and bounded, and non-essential: the tensed form could, at some cost to perspicuity, be written without it using a purely local composition operator $\circ$ (Steedman, 1985/1988; Jacobson, 1990), as $TP\backslash DP: pres \circ work$, or using the combinatory calculus of Curry and Feys (1958), as: $TP\backslash DP: \mathsf{B}\,pres\,work$.

[8] All logical forms are ruthlessly simplified for ease of reading. See Steedman, 2012 and section 4 on semantics of quantifier determiners like *a*. Elementary syntactic types like S and NP are also proxies for a finer-grain feature-based categorial type-system, such as that of Beavers (2004).

[9] This switch in subject-object dominance relations between derivation and predicate-argument structure has the effect of the syntactic WRAP operator used in other categorial frameworks (Bach, 1979, 1980; Jacobson, 1999), which in present terms is an exclusively lexical process.

follows from that fact. Agreement is reduced to simple matching of (possibly underspecified) features like 3*s*, with functors like *works*, $S\backslash NP_{3s}$, acting as "probes", where $\backslash NP_{3s}$, and the corresponding λ-bound variable y at lf again correspond to "uninterpretable" features. A contiguous *NP* and its lf value then act as a "goal", carrying "interpretable" features such as *3s* and *egon* (Adger, 2003; Radford, 2004). "Checking" then reduces to matching of feature-value pairs, either or both of which may be underspecified with respect to the other (Adger, 2010; Smith and Cormack, 2015). "Spell-out" of such constituents is immediate, since phonology and logical form are projected in lock-step from the lexical array by the derivation, constituting the "output levels" for the articulatory and inferential systems without further structural manipulation being necessary, or even possible under the adjacency assumption (3) (cf. Brody, 2002:22, Epstein and Seely, 2006:12; Stroik, 2009:14).[10]

A further important distinction between standard Minimalism and the present theory is that logical form is the only representational level it includes. The derivation that delivers the lf/pf pair is non-representational, in the sense that no rule can be conditioned on derivation structure, and syntactic derivation itself is entirely blind to lf representation. All rules of merger are conditioned only on categorial type, rather than derivation structure or lf. Phenomena that have in standard Minimalist approaches been held to be dependent on derivation structure, such as Superiority, Island effects, and Late Merge, are captured below in lexical specificational locality, projected onto derived categories by such rules, considered next.

2.2 Merge

The only rules required for the derivations in the preceding section are the two simplest kinds of directionally-specified contiguous merge, which can be written as follows:

(11) *Contiguous Merge I: The Application Rules:*

a. Forward Application

$$X/_{\star}Y : f \quad Y : a \quad \Rightarrow \quad X : f\,a \qquad (>)$$

b. Backward Application

$$Y : a \quad X\backslash_{\star}Y : f \quad \Rightarrow \quad X : f\,a \qquad (<)$$

X and Y are variables over categories like S, NP_{3s}, and $S\backslash NP_{3s}$, while f and a are variables over the corresponding logical forms. It will be convenient to refer to the functor X/Y or $X\backslash Y$ that determines the result-type X in such rules as the "governing category" and Y (etc.) as the "dependent category". The type $\star$ on the slashes in these rules means that only categories whose own slash-type is compatible can combine by this rule. (This detail can be ignored for now, since the unadorned slashes we have seen so far are compatible with any rule, including these.)[11]

The above rules reflect the following two generalization (Steedman, 1987), which hold of all versions of merger proposed here:

(12) *The Principle of Consistency:* All rules linearize their inputs consistently with the directionality specified in the governing category.

---

[10] The logical forms shown in (9) and (10) for English and Welsh are of course both simplified and language-specific. That is, elements like *past* should be thought of as themselves corresponding to more complex substructures of logical expressions of the general kind discussed by Cinque (2013) as "cartographic", which should be defined via lexical logical form, without mediation by movement or devices like unfilled heads (Brody, 1995).

[11] Typed categorial slashes were first proposed by Hepple (1990). The particular type-system used here is due to Baldridge (2002).

(13) *The Principle of Inheritance:* Any category that appears in an input that also appears in the output of a rule must be feature-identical in both, including its slash-features, if any.

Rules like the following are thereby disallowed, as indicated by the non-reduction symbol $\not\Rightarrow$:

(14) a. $Y : a \quad X/Y : f \quad \not\Rightarrow \quad X : f\,a$
b. $(X/Y)/W : f \quad Y : a \quad \not\Rightarrow \quad X/W : \lambda w.f\,w\,a$
c. $X_i/Y : f \quad Y : a \quad \not\Rightarrow \quad X_j : f\,a$

Consistency and Inheritance are corollaries of the Categorial and Adjacency assumptions (2): and (3), which also entail Inclusiveness, Extension, and No-tampering Conditions.

It is natural to ask if there are any other rules for combining categories which are similarly transparent with respect to syntax and semantics.

The simplest thing that you can do with a function, other than applying it, is to compose it with another function, to yield a new function. In the linearized notation used above for application-merge, rules of composition-merge can be written as follows:[12]

(15) *Contiguous Merge IIa: The Composition Rules (*$\mathbf{B}$*):*
a. Forward Composition:
$X/_{\diamond}Y : f \quad Y/Z : g \quad \Rightarrow \quad X/Z : \lambda z.f(g\,z)$ $\quad (>\mathbf{B}\ )$
b. Backward Composition:
$Y\backslash Z : g \quad X\backslash_{\diamond}Y : f \quad \Rightarrow \quad X\backslash Z : \lambda z.f(g\,z)$ $\quad (<\mathbf{B}\ )$
c. Forward Crossing Composition:
$X/_{\times}Y : f \quad Y\backslash Z : g \quad \Rightarrow \quad X\backslash Z : \lambda z.f(g\,z)$ $\quad (>\mathbf{B}_{\times})$
d. Backward Crossing Composition:
$Y/Z : g \quad X\backslash_{\times}Y : f \quad \Rightarrow \quad X/Z : \lambda z.f(g\,z)$ $\quad (<\mathbf{B}_{\times})$

The Principles of Consistency (12) and Inheritance (13) apply to the composition rules (15): rules like the following are thereby disallowed:

(16) $X/_{\diamond}Y : f \quad Y/Z : g \quad \not\Rightarrow \quad X\backslash Z : \lambda z.f(g\,z)$ $\quad (*{>}\mathbf{B})$

The composition rules (15) resemble function application merge (11) in having the effect of checking or canceling Y. Like application, they are syntactically transparent to semantic functional composition at the level of their uniform logical form, $\lambda z.f(g\,z)$. The rules have the effect of abstracting over z via contiguous merger, projecting the binding of z from g over f.[13]

The $\diamond$ and $\times$ slash-types on the governing functor X/Y or $X\backslash Y$ in the composition rules (15), like the $\star$ slash-type in the application rules (11), mean that only categories whose own slash is lexically specified as compatible can compose with dependent functor categories Y/Z or $Y\backslash Z$ via these rules. As in the case of application merger, we can ignore these details for now, since the plain slash categories seen so far are defined as combining by any rule, including these as well as application: However, they will become important when we need to distinguish possibilities for scrambling among languages like English and German.

When they apply in syntactic derivations, rules of function composition have the effect of generalizing the standard linguistic notion of constituency (Wells, 1947). For example, they

---

[12] The index $\mathbf{B}$ reflects the combinatory notation for function composition of Curry and Feys (1958), which could at some cost in readability be used in the logical forms, replacing $\lambda z.f(g\,z)$ by variable-free $\mathbf{B}\,f\,g$ (cf. note 7).

[13] Again, the use of λ binding in the lf notation is non-essential: the abstractions in (15) could be defined in variable-free terms using a purely local composition operator $\circ$ or Curry's $\mathbf{B}$ at lf (cf. nn.7, 12.)

allow an auxiliary $(S\backslash NP)/VP$ to compose with a transitive verb VP/NP in advance of combination of the latter with an object to form a constituent of the same type as a transitive verb:[14]

(17)
$$\dfrac{\dfrac{\text{will}}{(S\backslash NP)/VP : \lambda p\lambda y.will(py)} \quad \dfrac{\text{eat}}{VP/NP : \lambda x\lambda y.eat\,x\,y}}{(S\backslash NP)/NP : \lambda x\lambda y.will(eat\,x\,y)}{>}\mathbf{B}$$

We will later justify more fully the non-standard view of constituency that is implicit in the above, merely noting here that, on the assumption (to be elaborated below) that coordination is an operation that applies to constituents of like-type, the fact that *will eat* can coordinate with/be substituted by lexical transitives in sentences like the following is provisional evidence for the soundness of the rule illustrated above:[15]

(18) a. I cooked$_i$ and *will eat*$_i$ a fish$_i$.
b. a fish that$_i$ I cooked$_i$ and *will eat*$_i$

The combinatory Principles of Consistency, and Inheritance, together with the Inclusiveness, Extension, and No-tampering Minimalist Conditions should not be confused with substantive constraints, such as the Minimalist Shortest Move Condition. The latter is part of the definition of the Minimalist theory of grammar, whose inclusion is necessary to prevent overgeneralization by rules like MOVE, and to allow learnability by children. By contrast, the former are all formal properties constitutive of the Computation, which follow as theorems from the assumptions (2) and (3) limiting merger to pure functional operations including APPLICATION and COMPOSITION, and from the grounding of directional slash features in string position of section 2.1, and which characterize the learning mechanism itself.

They can collectively be summarized informally in the following generalization (Steedman, 2000b):

(19) *The Combinatory Projection Principle (CPP)*: Combinatory rules apply to contiguous categories (“Adjacency”), must respect the linearization specified in the slash direction for the governing category (“Consistency”), and must project unchanged onto the resulting category any further categorial, selectional, and linearization information specified in either the governing or the dependent category (“Inheritance”).

If we can compose a governing $X|Y$ into a unary dependent category $Y|Z$ (where “|” schematizes over “$/$” and “$\backslash$”), then since the lexicon includes binarized versions of multi-valent categories like transitive and ditransitive verbs, it is reasonable to also allow composition into dependent functors of the form $(Y|Z)|W$, again “canceling” Y to yield a result of the form $(X|Z)|W$, where “|” is the same in the input and result.

The full set of second-level composition rules is the following, entirely parallel to (15):[16]

---

[14] In Steedman (2002, 2018), I suggested that the possibility of recruiting composition and substitution combinators to language may have arisen from a primordial use in planning, as the operator for sequencing actions, defined as functions from situations to situations.

[15] Coordination is of course one of the traditional tests for constituency. We will return to this point in the Conclusion.

[16] In Steedman, 2014a, I suggested that the possibility of recruiting level-2 composition rules to language was due to its prior development for planning with actions involving tools or other agents.

(20) *Contiguous Merge IIb: The Second-level Composition Rules ($\mathbf{B}^2$):*

a. Forward Level-2 Composition:

$$X/_{\diamond}Y : f \quad (Y/Z)|W : g \quad \Rightarrow \quad (X/Z)|W : \lambda w\lambda z.f(gwz) \qquad ({>}\mathbf{B}^2)$$

b. Backward Level-2 Composition:

$$(Y\backslash Z)|W : g \quad X\backslash_{\diamond}Y : f \quad \Rightarrow \quad (X\backslash Z)|W : \lambda w\lambda z.f(gwz) \qquad ({<}\mathbf{B}^2)$$

c. Forward Crossing Level-2 Composition:

$$X/_{\times}Y : f \quad (Y\backslash Z)|W : g \quad \Rightarrow \quad (X\backslash Z)|W : \lambda w\lambda z.f(gwz) \qquad ({>}\mathbf{B}^2_{\times})$$

d. Backward Crossing Level-2 Composition:

$$(Y/Z)|W : g \quad X\backslash_{\times}Y : f \quad \Rightarrow \quad (X/Z)|W : \lambda w\lambda z.f(gwz) \qquad ({<}\mathbf{B}^2_{\times})$$

These rules conform to the CPP principles of Adjacency, Consistency, and Inheritance, so that alternatives like the following are universally disallowed, similarly to (16):

(21) $$X/Y : f \quad (Y/Z)/W : g \quad \not\Rightarrow \quad (X/Z)\backslash W : \lambda w\lambda z.f(gwz) \qquad ({*}{>}\mathbf{B}^2)$$

In syntactic derivations, such rules allow compositions like the following

(22) $$\begin{array}{cc} \text{may} & \text{sell} \\ \hline (S\backslash NP)/VP : \lambda p\lambda y.may(py) & (VP/PP)/NP : \lambda w\lambda x\lambda y.sell\,xwy \\ \hline \multicolumn{2}{c}{((S\backslash NP)/PP)/NP : \lambda w\lambda x\lambda y.may(sell\,xwy)} \end{array} \; {>}\mathbf{B}^2$$

Such level-2 composition mergers will be seen later to play the same role as level 1 composition (18) in coordinate sentences and relative clauses like the following:

(23) a. I [may sell]$_{((S\backslash NP)/PP)/NP}$ and [might give]$_{((S\backslash NP)/PP)/NP}$ my big pink Cadillac to my very best friend.

b. A car that I [may sell]$_{((S\backslash NP)/PP)/NP}$ and [might give]$_{((S\backslash NP)/PP)/NP}$ to my best friend

Since X in the governing category $X|Y$ and the result $(X|Z)|W$ of the second-level combinatory rules (20) can, as in all combinatory rules, match a function of n arguments (such as $S\backslash NP$ in (22)), so that $X|Y$ itself has valency $n+1$, and since the result $(X|Z)|W$ has one more argument than $X|Y$, making its valency $n+2$, and since that result can act as the governing category in a further application of level-2 composition, it should be obvious that repeated application of the second-level rules (20) can "grow" categories of unboundedly high valency. The set of non-terminal category types is therefore unbounded. Since it is a defining property of context-free grammars that the set of non-terminals is bounded, the theory of grammar we are dealing with is clearly non-context-free.[17]

A further class of combinatory rules corresponding to linearized versions of the "substitution" combinator **S** was originally proposed by Szabolcsi (1983), and is implicated in the analysis of parasitic gaps (Szabolcsi, 1983/1992; Steedman, 1987, 1996). These rules can be ignored for present purposes, but also constitute cases of contiguous merger, subject to the CPP conditions of Adjacency, Consistency, and Inheritance.

This completes the syntactic component of the grammar. No other rules are allowed. In these rules, agreement and "transfer" to the articulatory and inference systems, or "spell-out", are en-

---

[17] We exemplify this power in section 4.2 below. The class of languages it allows is among the least expressive non-context-free class that is known—not merely "mildly context sensitive" or LCFRS (Joshi, 1985), but "near- context-free", like Tree Adjoining Grammar (TAG), to which it is only weakly equivalent (Weir, 1988; Joshi et al., 1991; Rogers, 2003; Koller and Kuhlmann, 2009; Stanojević and Steedman, 2021), rather than strongly equivalent in the original sense of that term (Miller, 1999). (The latter is not the same as that in Schiffer and Maletti, 2021).

tirely synchronous with merger. Despite the involvement of three distinct varieties of contiguous merger—application, composition, and the substitution variety omitted here, such an architecture is therefore potentially simpler than is standard in Minimalist approaches (Jackendoff, 1997; Chomsky, 2001/2004). It remains to be shown how this simplicity can be maintained in the face of the major constructions that engender discontinuous dependency, which are considered in the remainder of this paper.

3 Raising and Composition

This section analyses raising and the related *there*-insertion construction as the first example of the generalization argued for here, namely that all instances of long-range dependency are mediated by composition of the same categories that establish local dependencies, using only the strictly contiguous combinatory rules of merger that were presented in the preceding section. In particular, *there*-insertion specifically requires a rule of composition merger (15a), and offers support a priori for the unorthodox notion of constituency engendered by the composition rules.

The varieties of discontinuity between predicates and their arguments that are bounded to a single tensed domain, such as the following are traditionally regarded under generative approaches as falling into two groups (Radford, 2004:268-274), exemplified by the following:

(24) a. *John* seems *to sleep*.
b. *John* tries *to sleep*.

The first, exemplified by (24a), consists of the “raising” constructions, which all transformational theories have described in terms of movement of subjects like *John* from infinitival predicates like *to sleep* to the specifier position of Tense, which in present lexicalized terms is simply the subject argument of *seemed*.

The second, exemplified by (24b), consists of the “(obligatory) control” constructions, which most recent generative accounts have viewed as mediated by an anaphoric PRO subject of non-finite predicates like “(to) sleep” (Chomsky, 1981; Chierchia, 1984; Landau, 2001, 2015, 2021), obligatorily bound to matrix arguments such as *John* by a variety of mechanisms.[18]

Both constructions can be analyzed by extending the set of lexical types considered so far to include certain second-order functions taking functions—more exactly, VP predicates or properties of semantic type $e \rightarrow t$ as their arguments (Chierchia, 1984). Here we will only consider control insofar as it contrasts with raising.

In present terms, raising and control must be expressed lexically, as in the following exemplars, where the syntactic type XP schematizes over predicative $PP, AP, NP, VP_{ing}, VP_{pss}$ (excluding VP, VP_{to}, VP_{en})—roughly, the attributive NP modifiers that appear elsewhere as $NP\backslash NP$:[19]

---

[18] Others, including Postal (1974), Lasnik (2001), Hornstein (1999, 2001), Boeckx et al. (2010), and Johnson (2020), have attributed the relation, like raising, to movement.

[19] Although we largely ignore control in the examples below, it may also be helpful in connection with *there*-insertion to note that passivization has the effect of mapping agentive raising-to-object and object-control verbs respectively into raising-to-subject and subject-control categories:

(i) *Passive of Raising-to-object:*
(be) believed := $VP_{psv}/VP_{to} : \lambda p \lambda x.believe(px)\,one$
Passive of Object-control:
(be) persuaded := $VP_{psv,+a}/VP_{to} : \lambda p \lambda x.persuade(px)\,x\,one$

(25) *Non-raising:*

think $:=$ $VP_{+a}/S : \lambda s \lambda y.thinks\,s\,y$

Raising-to-subject/Copula:

be $:=$ $VP/XP : \lambda p \lambda y.p\,y$

seem $:=$ $VP/VP_{to} : \lambda p \lambda y.seem(p\,y)$

likely $:=$ $AP/VP_{to} : \lambda p \lambda y.probable(p\,y)$

to $:=$ $VP_{to}/VP : \lambda p \lambda y.p\,y$

Raising-to-object:

believe $:=$ $(VP/VP_{to})/NP : \lambda x \lambda p \lambda y.believe(p\,x)\,y$

Subject-control:

try $:=$ $VP_{+a}/VP_{to} : \lambda p \lambda y.try(p\,y)\,y$

Object-control:

persuade $:=$ $(VP_{+a}/VP_{to})/NP : \lambda x \lambda p \lambda y.persuade(p\,x)\,x\,y$

Adjunct-control:

without $:=$ $(VP\backslash VP)/VP_{ing} : \lambda p \lambda q \lambda y.\neg p\,y \wedge q\,y$

The raising and copular categories listed above have the effect of composing their lf content (if any) with that of their complement predicate, and are akin to Jacobson's 1990 analysis, which restricts related categories to syntactic combination by the forward composition rule (15a), although function composition here is lexicalized at lf, as in Dowty, 1978 and Sag, 1982, as shown in the following derivations, rather than syntactic.

(26)
$$\begin{array}{cccc}
\text{Marcel} & \text{seems} & \text{to} & \text{sleep} \\
\hline
NP : marcel & (S\backslash NP)/VP_{to} : \lambda p \lambda y.pres(seem(p\,y)) & VP_{to}/VP : \lambda p \lambda y.p\,y & VP : \lambda y.sleep\,y \\
\cline{3-4}
 & & \multicolumn{2}{c}{VP_{to} : \lambda y.sleep\,y \quad >} \\
\cline{2-4}
 & \multicolumn{3}{c}{S\backslash NP : \lambda y.pres(seem(sleep\,y)) \quad >} \\
\cline{1-3}
\multicolumn{3}{c}{S : pres(seem(sleep\,marcel)) \quad <} &
\end{array}$$

(27)
$$\begin{array}{cccc}
\text{Marcel} & \text{tries} & \text{to} & \text{sleep} \\
\hline
NP : marcel & (S_{+a}\backslash NP)/VP_{to} : \lambda p \lambda y.pres(try(p\,y)\,y) & VP_{to}/VP : \lambda p \lambda y.p\,y & VP : \lambda y.sleep\,y \\
\cline{3-4}
 & & \multicolumn{2}{c}{VP_{to} : \lambda y.sleep\,y \quad >} \\
\cline{2-4}
 & \multicolumn{3}{c}{S_{+a}\backslash NP : \lambda y.pres(try(sleep\,y)\,y) \quad >} \\
\cline{1-3}
\multicolumn{3}{c}{S_{+a} : pres(try(sleep\,marcel)\,marcel) \quad <} &
\end{array}$$

The distinction between raising and control verbs at the level of logical form is that, in raising in (26), the bound variable y occurs only once as a subject variable in the predicate-argument structure, where it is bound by its λ-binder, whereas in control (27), the bound variable occurs twice at that level, once as subject or object controller and once as controllee, forming a potentially unbounded cascade, a distinction parallel to that between A-movement and the PRO mechanism in the Government-Binding theory. In both cases, the traditional c-command relation between raiser/controler and complement needed to support scopal operator binding is only established at the level of lf:[20]

---

[20] Landau (2015), following Williams (1994), draws a number of finer semantic distinctions among subject- and object-control verbs that are passed over here, including a distinction between "predicative" ones like *manage* and *begin*, and "attitudinal" ones like *hope*, *persuade*, and *tell*, with distinctions in factivity and obligatoriness or otherwise of "de se" readings (Lewis, 1979a; Chierchia, 1989) based on scope of intensional operators at the level of logical form. The lexical

The categories in (25) assume in addition that raising and copular verbs and adjuncts like *likely* are, unlike most other verbs, lexically underspecified as to agentivity on their their complement and their result via a feature variable a, whose value $\pm$ they merely inherit from their complement (which of course may also be underspecified), and which is left unmarked in the notation, to avoid clutter. By contrast, the control verbs are $+a$(gentive). The feature a does work reminiscent of the *there* feature in Sag (1982):444 in limiting overgeneration, but plays no part in mediating discontinuity, including long-range agreement. While to keep the notation readable, we suppress such details elsewhere in the paper, its underspecification in the raising/copular categories will be relevant to the analysis of *there*-insertion below.

3.1 Raising

This section shows that the A-movement variety of internal merge involved in raising is captured by the rules of contiguous type-dependent merger set out in section 2.2. In order to extend the analysis to *there*-insertion in the next section, we focus on raising out of the copular VP, as in:

(28)

$$
\begin{array}{ccccl}
\text{Marcel} & \text{seems} & \text{to be} & \text{asleep.} & \\
\hline
NP_{3s} & (S\backslash NP_{3s})/VP_{to} & VP_{to}/XP & AP & \\
:marcel & :\lambda p\lambda y.pres(seem(py)) & :\lambda p\lambda y.py & :\lambda y.asleep,y & \\
\cline{3-4}
 & & \multicolumn{2}{c}{VP_{to}:\lambda y.asleep\,y} & > \\
\cline{2-4}
 & \multicolumn{3}{c}{S\backslash NP_{3s}:\lambda y.pres(seem(asleep\,y))} & > \\
\cline{1-4}
\multicolumn{4}{c}{S:pres(seem(asleep\,marcel))} & <
\end{array}
$$

Unlike control in (27), raising across serial raising verbs does not create a cascade of multiple copies of the subject, which remains in situ at lf (cf. Epstein and Seely, 2006:§2.4):

(29)

$$
\begin{array}{cccccl}
\text{Marcel} & \text{seems} & \text{to appear} \ldots & \text{to be} & \text{dreaming} & \\
\hline
NP_{3s} & (S\backslash NP_{3s})/VP_{to} & VP_{to}/VP_{to} & VP_{to}/XP & VP_{ing} & \\
:marcel & :\lambda p\lambda y.pres(seem(py)) & :\lambda p\lambda y.appear(py) & :\lambda p\lambda y.py & :\lambda y.prog(dream\,y) & \\
\cline{4-5}
 & & & \multicolumn{2}{c}{VP_{to}:\lambda y.prog(dream\,y)} & > \\
\cline{3-5}
 & & \multicolumn{3}{c}{VP_{to}:\lambda y.appear(prog(dream\,y))} & > \\
\cline{2-5}
 & \multicolumn{4}{c}{S_{to}\backslash NP_{3s}:\lambda y.pres(seem(appear(prog(dream\,y))))} & > \\
\cline{1-5}
\multicolumn{5}{c}{S:pres(seem(appear(prog(dream\,marcel))))} & <
\end{array}
$$

Raising verbs can also merge by successive composition, when they form a category of the same type as a raising verb, as in the following alternative derivation for (29):

(30)

$$
\begin{array}{cccccl}
\text{Marcel} & \text{seems} & \text{to appear} \ldots & \text{to be} & \text{dreaming} & \\
\hline
NP_{3s} & (S\backslash NP_{3s})/VP_{to} & VP_{to}/VP_{to} & VP_{to}/XP & VP_{ing} & \\
:marcel & :\lambda p\lambda y.pres(seem(py)) & :\lambda p\lambda y.appear(py) & :\lambda p\lambda y.py & :\lambda y.prog(dream\,y) & \\
\cline{2-3}
 & \multicolumn{2}{c}{(S\backslash NP_{3s})/VP_{to}:\lambda p\lambda y.pres(seem(appear(py)))} & & & >\mathbf{B} \\
\cline{2-4}
 & \multicolumn{3}{c}{(S\backslash NP_{3s})/XP:\lambda p\lambda y.pres(seem(appear(py)))} & & >\mathbf{B} \\
\cline{2-5}
 & \multicolumn{4}{c}{S_{to}\backslash NP_{3s}:\lambda y.pres(seem(appear(prog(dream\,y))))} & > \\
\cline{1-5}
\multicolumn{5}{c}{S:pres(seem(appear(prog(dream\,marcel))))} & <
\end{array}
$$

---

logical forms shown here are compatible with such finer distinctions, but are underspecified for the present purpose. Landau, 2021: 20 accounts for adjunct control by lexical specification, as in (25) above.

More specifically, raising categories can compose with the copula to yield a category of the same type as the copula, $(S\backslash NP_{agr})/XP$. The existential *there*-insertion construction, considered next, applies across just such sequences of raising and copular categories.[21]

3.2 There-*insertion*

To understand the existential *there* construction, it is important to recall Carlson's (1977) distinction between "stage-level" predicates, which denote "fluents" or transient properties like *at the bottom of our garden*, which are bounded in temporal extent, and "individual-level" predicates, which denote intrinsic properties with unspecified temporal extent, like *good*:[22]

The existential *there*-insertion construction exhibits discontiguous subject agreement across potentially unbounded sequences of raising and copular categories, and is only compatible with stage-level (transient) predicates:[23]

(31) a. Fairies are/seem to be good/at the bottom of our garden
b. There are/seem to be fairies *good/at the bottom of our garden.

The NP complement in the *there*-construction also has to be indefinite:

(32) a. There are fairies/some fairies/many fairies/no fairies at the bottom of our garden.
b. There is *the fairy Paribanou/*every fairy/*it at the bottom of our garden.

I follow Bolinger (1977), Rando and Napoli (1978), Abbott (1993), and Huddleston and Pullum (2002:1392-1403) in assuming both constraints to be essentially pragmatic in origin, related to discourse "newness" (Prince, 1981):

Williams (1984) and Jacobson (1990), as well as the related G/HPSG feature-passing approaches of Gazdar et al. (1985) and Levine (2017:186), and the TAG-based Minimalist approach of Frank (2002:113), account for the existential *there*-construction by assigning the verbs involved an additional more specialized lexical category specifying an expletive NP_{there} subject.

The alternative approach followed here takes advantage of the fact that serial raising verbs can compose to yield a non-standard constituent with the same category as the copula, as in (30), by making *there* the head of the construction, assigning it the lexical categories shown in figure 1a,c.[24]

---

[21] This observation is implicit in the related "raising as function composition" analysis of Jacobson, 1990, and is related to the treatment of German "long passives" by Keine and Bhatt (2016).

[22] While in English the stage/individual-level distinction is not marked in morpho-syntax, it is in other languages. Kratzer (1988/1995) and Diesing (1992) ground the distinction in the semantics, where stage-level predicates include a spatio-temporally locative Davidsonian lf event-variable, which individual-level predicates lack. The present paper passes over this distinction in logical forms, to simplify. Like other aspectual distinctions, it is labile: the predicates usually found to be stage-level can in contexts requiring individual level predicates be "coerced" to the latter type, and vice versa.

[23] The construction is topic-establishing, and is often to be found in the opening lines of Edwardian dramatic monologues ("There's a one-eyed yellow idol to the north of Kathmandu ..."). We pass over a further class of verbs like *arrive*, *arise*, *appear*, etc. which can occur in the *there*-construction with a similar topic-establishing effect when they are predicated of indefinite and stage-level complements, and appear to bear the category of the copula:

(i) a. There appeared a tall ship on the horizon.
b. There arrived a train in the station.
c. There hung a shotgun upon the wall.

See Levin (1993:section 6.1), Hale and Keyser (2002), and Deal (2009) for extensive discussion.

[24] Multiple occurrences of the variable *arg* denote the same underspecified value. Discourse-semantic details of how the resulting category is represents indefinite subjects and stage-level predicates are passed over. A further subject-

a. Subject *there*:

$$\text{there} \quad := \quad ((S/XP_{+stg})/NP_{agr,-def})/((S_{-a}\backslash NP_{agr})/XP) : \lambda c\lambda y\lambda p.c(p\,y)$$

b.

$$\begin{array}{cccc}
\text{There} & \text{seem to appear to be} & \text{fairies} & \text{at the bottom of our garden.} \\
\cline{1-1}\cline{2-2}\cline{3-3}\cline{4-4}
((S/XP_{+stg})/NP_{agr,-def})/((S_{-a}\backslash NP_{agr})/XP) & (S\backslash NP_{3p})/XP \;{}^{>\mathbf{B}^*} & NP_{3p} & PP \\
:\lambda c\lambda y\lambda p.c(p\,y) & :\lambda p\lambda y.pres(seem(appear(p\,y))) & :fairies & :\lambda y.at(bottom\,garden)\,y \\
\cline{1-2}
\multicolumn{2}{c}{(S/XP_{+stg})/NP_{3p,-def} : \lambda y\lambda p.pres(seem(appear(p\,y))) \quad >} & & \\
\cline{1-3}
\multicolumn{3}{c}{S/XP_{+stg} : \lambda p.pres(seem(appear((p\,fairies)))) \quad >} & \\
\cline{1-4}
\multicolumn{4}{c}{S : pres(seem(appear(at(bottom\,garden)\,fairies))) \quad >}
\end{array}$$

c. Object *there*:

$$\text{there} \quad := \quad ((((S\backslash NP)/XP_{+stg})/NP_{-def})/(VP_{to}/XP))\backslash(((S\backslash NP)/VP_{to,-a})/NP) : \lambda b\lambda x\lambda c\lambda y.b(c\,x)\,y$$

d.

$$\begin{array}{cccccc}
\text{I} & \text{believed} & \text{there} & \text{to be} & \text{fairies} & \text{at the bottom of our garden} \\
\cline{1-1}\cline{2-2}\cline{3-3}\cline{4-4}\cline{5-5}\cline{6-6}
NP & ((S\backslash NP)/VP_{to})/NP & ((((S\backslash NP)/XP_{+stg})/NP_{-def})/(VP_{to,-a}/XP))\backslash(((S\backslash NP)/VP_{to})/NP) & VP_{to}/XP & NP & PP \\
\cline{2-3}
 & & \multicolumn{1}{c}{(((S\backslash NP)/XP_{+stg})/NP_{-def})/(VP_{to,-a}/XP) \quad <} & & & \\
\cline{3-4}
 & & \multicolumn{2}{c}{((S\backslash NP)/XP_{+stg})/NP_{-def} \quad >} & & \\
\cline{3-5}
 & & \multicolumn{3}{c}{(S\backslash NP)/XP_{+stg} \quad >} & \\
\cline{3-6}
 & & \multicolumn{4}{c}{S\backslash NP \quad >} \\
\cline{1-3}
\multicolumn{3}{c}{S \quad <} & & &
\end{array}$$

Figure 1: *There* insertion

The category for *there* in figure 1a, applies to a constituent to its right bearing the type of an SVX copula compatible with the non-agentivity feature-value $-a$—that is, $-a$ itself or the unmarked a of the raising and copular verbs. Composite raising/copular verbs of type $(S\backslash NP_{3p})/XP$ can be composed of unboundedly many raising elements, as in (30). What existential *there* does is simply to map such compound SVX copulæ as *are*, *seem to be*, *seem likely to be*, *seem to be believed to be* etc., onto the corresponding VSX verbal category *there are*, *there seem to be*, etc., further specifying the subject S and complement X as respectively indefinite and stage-level via the feature-values $-def$ and $+stg$, on which we assume most NPs like *fairies* and XPs like *at the bottom of our garden* are underspecified or ambiguous. They are therefore shown unmarked, with undifferentiated semantics for simplicity of presentation, in the *there*-insertion derivation shown in figure 1b.

The categories for *there* in figure 1a,c reject control verbs because of their positive $+a$(gentivity). For example:

(33)
$$\begin{array}{cccc} \ast\text{There} & \text{try to be} & \text{fairies} & \text{at the bottom of our garden.} \\ \hline ((S/XP_{+stg})/NP_{agr,-def})/((S_{-a}\backslash NP_{agr})/XP) & (S_{+a}\backslash NP_{3p})/XP \; {}^{>\mathbf{B}^{*}} & NP_{3p} & PP \\ \hline \multicolumn{2}{c}{***} \; {}^{***} & & \end{array}$$

The slightly more complex category for expletive *there* in figure 1c inverts the rightward arguments of raising-to-object verbs like *believe*, to yield a category much like that in figure 1a, looking for a non-agentive *to*-infinitival copular category such as *to be/to be certain to be believed to be* of type VP_{to}/XP, an (indefinite) NP, and a (stage-level) predicate. as in the derivation in figure 1d.[25]

The lexical categories of *there* in figure 1a and c prevent its application to *seem* or *are believed*, because they bear non-copular raising categorial types like $(S\backslash NP_{3p})/VP_{to}$ or $(S\backslash NP)/S$, rather than the copular categories specified by *there*, such as $(S\backslash NP_{3p})/XP$ and VP_{to}/XP. We therefore avoid the need for filters such as the Θ-criterion or "defective intervention effects" to avoid overgeneration of examples like the following (Chomsky, 1981, 2000:129): cf. Frampton and Gutmann, 2002; Stroik, 2009):

(34) a. *There seem fairies to be at the bottom of our garden.
b. *There are believed fairies are at the bottom of our garden

As noted earlier, verb sequences with the same type $(S\backslash NP)/XP$ as the copula that include agentive verbs such as *try* or *say* inherit the specified $+a$(gentive) feature of those verbs, via unification with the underspecified feature of raising and copular verbs, and are incompatible with the $-a$ required by*there*, figure 1a,c, for the same reason as (33):

(35) a. *There [seem to try to be]$_{(S_{+a}\backslash NP)/XP}$ fairies at the bottom of our garden.
b. *There [seem to say old bicycles are]$_{(S_{+a}\backslash NP)/XP}$ fairies at the bottom of our garden.

In summary, the long-distance agreement (LDA) discontinuity induced by raising and *there*-insertion can be captured by lexicalizing the verbs (etc.) involved as functions over predicates,

---

inversion category for *there*, related to figure 1c but looking to the left for the inverting copula $((S_{inv}/XP)/NP_{agr})$ as its first argument, is also needed to support questions like:

(i) a. Are there (believed to be) fairies at the bottom of our garden?
b. Where are there (likely to be) fairies?

We pass over it here, in the interests of brevity.

[25] Logical form and details of agreement are suppressed in figure 1, as analogous to earlier examples.

representing the raising discontinuity by λ-bound variables in lexical logical forms such as $\lambda p\lambda y.seem(p\,y)$, and projecting it onto larger domains by contiguous external merger of categories of the same type as copular verbs via the composition rule (15a).

4 Scrambling, Case, and Type-raising

As soon as we look beyond simple existential NPs such as proper names and indefinites, and in particular to universally quantified NPs, it becomes clear that arguments must themselves be second-order "type-raised" functors of the general form $\mathrm{X}/(\mathrm{X}\backslash NP)$ or $\mathrm{X}\backslash(\mathrm{X}/NP)$, where X is a variable ranging over lexical types S, $S\backslash NP$, $(S\backslash NP)/NP$, etc..

4.1 Type-raising as case

For example, for the universally-quantified subject of *Everything flows* to take scope over the tensed domain $S\backslash NP$ of its verb, it must syntactically bear the category of a functor over it, as in the following derivation:

(36)
$$\begin{array}{ccl}
\text{Everything} & \text{flows} & \\
\cline{1-1}\cline{2-2}
S/(S\backslash NP) & S\backslash NP & \\
:\lambda p.\forall y[thing\,y \Rightarrow p\,y] & :\lambda y.pres(flow\,y) & \\
\cline{1-2}
\multicolumn{2}{c}{S:\forall y[thing\,y \Rightarrow pres(flow\,y)]} & >
\end{array}$$

Because universals can freely coordinate with other NPs, this generalization must, under the earlier assumption that coordination applies over like-types, apply to all NPs (Montague, 1973):

(37) [Every woman]$_{S/(S\backslash NP)}$ and [Harry/at least one man]$_{S/(S\backslash NP)}$ [saw a movie]$S\backslash NP$.

We must make a parallel assumption of categorial raising over the transitive verb category $(S\backslash NP)/NP$ to allow objects to take scope over the matrix domain:[26]

(38)
$$\begin{array}{cccl}
\text{Someone} & \text{sees} & \text{everything} & \\
\cline{1-1}\cline{2-2}\cline{3-3}
S/(S\backslash NP) & (S\backslash NP)/NP & (S\backslash NP)\backslash((S\backslash NP)/NP) & \\
:\lambda p.p(some\,person) & :\lambda x\lambda y.pres(see\,x\,y) & :\lambda p\lambda y.\forall x[thing\,x \Rightarrow p\,x\,y] & \\
\cline{2-3}
 & \multicolumn{2}{c}{S\backslash NP:\lambda y.\forall x[thing\,x \Rightarrow pres(see\,x\,y)]} & < \\
\cline{1-3}
\multicolumn{3}{c}{S:\forall x[pres(see\,x(some\,person))]} & >
\end{array}$$

It should be obvious from the above examples that type-raising leaves all the derivations seen in the earlier sections essentially unchanged, apart from reversing the direction of function application. In particular, the above derivations still use rules of strictly external merger. However, it is now the raised argument categories that project the type S of the verbal head as the category or "label" of the result of the derivation. Inspection of the lexical logical forms in the derivation also shows that it is the raised arguments that pass their values to the lf by binding a second-order variable p, rather than the verb itself, with an effect that is reminiscent of "internal" merge, and

---

[26] It should be noted that such scope-taking is defined by the lexical lf of the universal, rather than by derivation. As in Steedman, 2012, non-universals are represented at lf in situ by under-specified Skolem terms like *some person*, to be bound at some point in the derivation by a mechanism whose details are irrelevant to the present purpose. If this happens early in the derivation at line 1 of (38), then the result is an unbound Skolem constant, analogous to a proper name. If it happens inside the lf scope of the universal, as in the last line, it is a Skolem function of x.

that we return to below. However, type-raised arguments also allow some additional word-orders which would not otherwise be possible, including some arising from function composition that have been attributed to "scrambling", considered in section 4.2 below.

The category $S/(S\backslash NP)$ of *everything* in (36) limits it to only combining as a subject. It is therefore equivalent in every way to a nominative NP in a language like Latin, despite lacking any morphology to distinguishing it from *everything* with category $(S\backslash NP)\backslash((S\backslash NP)/NP)$ in (38). The latter category is similarly restricted to combining as an object, equivalent to an accusative-cased NP in Latin. In fact, under present assumptions, English differs from Latin only that case is even more ambiguous (because not in general morphologically marked), and "structural" and "inherent", that is, defined semantically by the valency and logical form of the verb.

In view of this ambiguity, and the fact that it is resolved in the derivation by the verb it combines with, we can assume as a matter of grammar that NPs and other argument categories are schematized in the lexicon as in the following example for *someone*:

(39) someone := $NP^{\uparrow} : \lambda p \lambda \dot{y}.p\,(some\,person)\,\dot{y}$

—where $NP^{\uparrow}$ schematizes over NP categories raised over verbal (etc.) categories of the language, which are assumed to obey the Natural Serialization (NS) and Cross-Category Harmony (CCH) Principles observed by Vennemann and Harlow (1977); Hawkins (1982), and Dryer (1991), namely that arguments other than the subject are specified with directionality consistent with that of the object—that is, languages seem to be consistently VOX* or OX*V. In the logical form, p ranges as usual over the logical forms of such verbs (etc.); $\dot{y}$ is a (possibly empty), in principle unbounded, list of further arguments of p; and $\lambda\dot{y}$ is the corresponding sequence of λ-bindings. (Because the number of arguments is unbounded, it is possible under certain circumstances for this schema to apply to derived categories of types that are not actually exemplified in the lexicon itself, so long as they are consistent with NS/CCH—see discussion of figure 2b,c.)[27]

If arguments are all type-raised as $NP^{\uparrow}$ etc., then determiners etc. must be functors into that raised schema, of the form $NP^{\uparrow}/N$ etc. For example, the indefinite determiner category is:

(40) a/an :=$NP^{\uparrow}/N : \lambda n \lambda p \lambda \dot{y}.p(a\,n)\,\dot{y}$

—where n ranges over nominal properties; $a\,n$ abbreviates an under-specified Skolem term of type n as noted for (38) above (n.26).

The determiner *every* carries a similar set of cased syntactic categories, which can be schematized under the same conventions as (40) as:

(41) every := $NP^{\uparrow}/N : \lambda n \lambda p \lambda \dot{y}.\forall x[n x \rightarrow p x \dot{y}]$,

This category corresponds to a universal "Generalized Quantifier" determiner (Woods, 1968a,b; Montague, 1970; Lewis, 1970), in which the quantifier $\forall x$ takes scope at the level of lf over a potentially unbounded predication over x via the variable p, allowing the homomorphism of derivation and syntactic composition to be preserved without any involvement of "covert" quantifier movement in derivations like the following (Steedman, 2012).[28]

---

[27] For performance reasons, in practice, the valency is low—possibly as low as the maximum available to lexical verbs. This definition of type-raising does not exclude the possibility that, for example, a VOX* language may also systematically include OV categories, as appears to be the case for Mandarin (Li and Thompson, 1975; Man and Steedman, 2023).

[28] See note 26 on scope of indefinites. The Montagovian approach of Jacobson (2014) within a related Categorial framework provides an alternative.

(42)
$$\begin{array}{ccccc}
\text{A} & \text{woman} & \text{saw} & \text{every} & \text{cat}\\
\overline{(S/(S\backslash NP_{3s}))/N_{3s}} & \overline{N_{3s}} & \overline{(S\backslash NP_{agr})/NP} & \overline{((S\backslash NP_{agr})\backslash((S\backslash NP_{agr})/NP))/N_{3s}} & \overline{N_{3s}}\\
:\lambda n\lambda p.p(a\,n) & :woman & :\lambda x\lambda y.past(see\,x\,y) & :\lambda n\lambda p\lambda y.\forall x[n\,x\rightarrow p\,x\,y] & :cat\\
\multicolumn{2}{c}{\overline{\qquad\qquad\qquad\qquad}{}^{>}} & & \multicolumn{2}{c}{\overline{\qquad\qquad\qquad\qquad\qquad\qquad}{}^{>}}\\
\multicolumn{2}{c}{S/(S\backslash NP_{3s})} & & \multicolumn{2}{c}{(S\backslash NP_{agr})\backslash((S\backslash NP_{agr})/NP)}\\
\multicolumn{2}{c}{:\lambda p.p(a\,woman)} & & \multicolumn{2}{c}{:\lambda p\lambda y.\forall x[cat\,x\rightarrow p\,x\,y]}\\
 & & \multicolumn{3}{c}{\overline{\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad}{}^{<}}\\
 & & \multicolumn{3}{c}{S\backslash NP_{agr}:\lambda y.\forall x[cat\,x\rightarrow past(see\,x\,y)]}\\
\multicolumn{5}{c}{\overline{\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad}{}^{>}}\\
\multicolumn{5}{c}{S:\forall x[cat\,x\rightarrow past(see\,x(a\,woman))]}
\end{array}$$

As Woods, Montague, Lewis, Partee (1976), and Barwise and Cooper (1981) pointed out, the type-raised generalized quantifier determiner category (41) is the only way to get universal quantifiers to take wide-scope monotonically, that is, without quantifier-raising (QR) movement. However, type-raising is allowed here only as a morpho-lexical schema, not as a free combinatory rule of syntactic derivation like those in section 2.2. Derivational type-raising would increase expressive power.[29]

The assumption of raised categories therefore amounts to the claim that even a morphologically impoverished language like English has case in exactly the same sense as a case-morphology-rich language like Latin or Icelandic, disambiguating case "structurally", in relation to the category of the verb.[30]

4.2 Scrambling in Germanic

Most standard British and American dialects of English disfavor scrambling of NP arguments. Unlike Japanese and German cased arguments, English (and Welsh) case-raised $NP^{\uparrow}$s like *pizza* can be prevented from combining by crossing composition by morpho-lexically specifying their slash-type as $/_{\diamond\star}$ or $\backslash_{\diamond\star}$ in their lexical entries. The crossing composition rule (15d) can only apply to governing categories of the form $X\backslash_{\times}Y$ whose slash-type, unlike that assumed for *pizza*, is compatible with the feature $\times$, so the following is blocked:[31]

(43)
$$\begin{array}{cccc}
*\text{John} & \text{gave} & \text{pizza} & \text{a very close friend.}\\
\overline{S/_{\diamond\star}(S\backslash NP)} & \overline{((S\backslash NP)/NP)/NP} & \overline{(S\backslash NP)\backslash_{\diamond\star}((S\backslash NP)/NP)} & \overline{(S\backslash NP)\backslash_{\diamond\star}((S\backslash NP)/NP)}\\
 & \multicolumn{2}{c}{\overline{\qquad\qquad\qquad\qquad\qquad\qquad\qquad}{}^{*<\mathbf{B}_{\times}}} & \\
 & \multicolumn{2}{c}{***} &
\end{array}$$

However, English VP adjuncts like *yesterday* are unrestricted, and freely allow scrambling by composing by the backward crossing rule, as in the following example of "Heavy NP-shift":

---

[29] Derivational type-raising is implicit in the "Reprojection" analysis of quantifier scope-taking of Hornstein and Uriagereka (2002), although they make no reference to the literature on type-raising and generalized quantifiers.

[30] Cf. Vergnaud (1977/2006). In particular, there is nothing in the present theory to prevent a language resolving case structurally for subjects bearing "quirky" non-nominative morphological case agreement, as is notoriously the case for Icelandic. The fact that scrambling/free word-order is strongly correlated with morphologically explicit case clearly reflects a performance-related need to (somewhat) limit derivational ambiguity, rather than a universal of grammar.

[31] Such details of the categorial notation can often be ignored, and are often suppressed in examples. The example (43) is not as bad as its "*" suggests, and may be allowed in some dialects—see section 7.2 on Right Edge Restrictions.

(44)

$$\begin{array}{cccc}
\text{Harry} & \text{saw} & \text{yesterday} & \text{a very close friend.} \\
\cline{1-1}\cline{2-2}\cline{3-3}\cline{4-4}
S/_{\diamond\star}(S\backslash NP) & (S\backslash NP)/NP & (S\backslash NP)\backslash(S\backslash NP) & (S\backslash NP)\backslash_{\diamond\star}((S\backslash NP)/NP) \\
\cline{2-3}
 & \multicolumn{2}{c}{(S\backslash NP)/NP \quad {<}\mathbf{B}_{\times}} & \\
\cline{2-4}
 & \multicolumn{3}{c}{S\backslash NP \quad {<}} \\
\cline{1-4}
\multicolumn{4}{c}{S \quad {>}}
\end{array}$$

Under the assumption that case-related morpho-lexical type-raising of arguments is universal, its availability in verb-final Germanic constructions correctly allows them to scramble over each other further to the left. The question considered next concerns possible limitations on such scrambling. Wallenberg (2009, 2013) proposes the following generalization:

(45) *The Generalized Holmberg Condition (*GHC*):*
Scrambling and object shift cannot move elements leftward past a c-commanding head.

He notes that in the German example (46), the object *die Lebensmittel* ("the groceries") can scramble out of the VP past the adjunct, as in (47):

(46) Johann hat [auf dem Markt] [die Lebensmittel gekauft].
Johann has at the market the groceries bought.
"Johann bought the groceries at the market."

(47) Johann hat [die Lebensmittel] [auf dem Markt] gekauft.
Johann has the groceries at the market bought.
"Johann bought the groceries at the market."

However, *die Lebensmittel* cannot scramble further to the left past the main clause auxiliary *hat*, which c-commands that VP:

(48) *Johann [die Lebensmittel] hat [auf dem Markt] gekauft.
Johann the groceries has at the market bought.
"Johann bought the groceries at the market."

(47) is possible because the PP adjunct *auf dem Markt* VP/VP that intervenes between the scrambled object and the verb can compose with it to yield a category $VP\backslash NP$ adjacent to the case-raised object:[32]

(49)

$$\begin{array}{ccccc}
\text{Johann} & \text{hat} & \text{die Lebensmittel} & \text{auf dem Markt} & \text{gekauft} \\
\text{Johann} & \text{has} & \text{the groceries} & \text{at the market} & \text{bought.} \\
\cline{1-1}\cline{2-2}\cline{3-3}\cline{4-4}\cline{5-5}
S/(S\backslash NP) & (S_{main}\backslash NP_{nom})/VP & VP/(VP\backslash NP) & VP/VP & VP\backslash NP_{acc} \\
\cline{4-5}
 & & & \multicolumn{2}{c}{VP\backslash NP_{acc} \quad {>}\mathbf{B}_{\times}} \\
\cline{3-5}
 & & \multicolumn{3}{c}{VP \quad {>}} \\
\cline{2-5}
 & \multicolumn{4}{c}{S\backslash NP_{nom} \quad {>}} \\
\cline{1-5}
\multicolumn{5}{c}{S \quad {>}}
\end{array}$$

However, if the same object is to the left of the main verb, it is no longer adjacent to a category of the required type, even though the tensed verb can compose with the residue of scrambling:

---

[32] The same applies if the PP is taken to be a type-raised argument $VP/(VP\backslash PP)$.

(50)
$$\begin{array}{ccccc}
*\text{Johann} & \text{die Lebensmittel} & \text{hat} & \text{auf dem Markt} & \text{gekauft}\\
\text{Johann} & \text{the groceries} & \text{has} & \text{at the market} & \text{bought.}\\
\overline{S/(S\backslash NP)} & \overline{VP/(VP\backslash NP_{acc})} & \overline{(S_{main}\backslash NP_{nom})/VP} & \overline{VP/VP} & \overline{VP\backslash NP_{acc}}\\
 & & & \multicolumn{2}{c}{\overline{\qquad\qquad\qquad\qquad}{>}\mathbf{B}_{\times}}\\
 & & & \multicolumn{2}{c}{VP\backslash NP_{acc}}\\
 & & \multicolumn{3}{c}{\overline{\qquad\qquad\qquad\qquad\qquad\qquad}{>}\mathbf{B}_{\times}}\\
 & & \multicolumn{3}{c}{(S_{main}\backslash NP_{nom})\backslash NP_{acc}}\\
\multicolumn{5}{c}{\overline{\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad}{*}}\\
\multicolumn{5}{c}{***}
\end{array}$$

There is no German morpholexically cased accusative category that can combine with *hat auf dem Markt gekauft*, since $(S_{main}\backslash NP_{nom})\backslash NP_{acc}$ is not a V2 category, and therefore not a legal German main-clause transitive. (The Germanic subordinate clause transitive category, which accusatives are also raised over, is $(S_{sub}\backslash NP_{nom})\backslash NP_{acc}$, while the main clause topicalized accusative is $S_t/(S_{inv}\backslash NP_{acc})$—cf. (58) below.).

In fact, it is only because arguments are lexically type-raised, and ground categories like NP_{acc} are excluded, that examples like (48) can in present terms be blocked, consistent with Wallenberg's generalization (45). The effect of type-raising arguments is to define the "domain of locality" for scrambling to be the argument domain of the verb they combine with.

For example, the application-only derivation shown in figure 2a for the class of Swiss German object-control verbs that take bare infinitival complements can be assumed to represent the unscrambled order (Shieber, 1985), resembling the corresponding English sentence, except for objects preceding their head: However, because Swiss-German object control verbs mix directionality, taking their object to the left and their infinitival complement to the right, composition has the effect of extending the verbal domain of locality, correctly predicting the fact that embedded objects can scramble leftward "past" tense, as in the derivation shown in figure 2b. The crucial first step in this derivation, involving the second-level composition rule (20c), indexed $>\mathbf{B}^2_{\times}$, can iterate unboundedly over additional object-control verbs, allowing embedded objects like *em Hans* and *es huus* to scramble to the left past the tensed verb, leading Wallenberg to replace (45) by a more general Principle of "Conservation of C-command" (CoCC). The residue of scrambling bears the category $(((S'\backslash NP_{nom})\backslash NP_{acc})\backslash NP_{dat})\backslash NP_{acc}$, which does not seem to be a lexical type, so the definition at the beginning of section 4.1 of type-raising as an unbounded lexical schema obeying the principle of VOX*/OX*V order is essential in allowing *es huus* to scramble[33]

Such scrambling or partial free-order of arguments in Germanic and other languages presents a problem for all theories of grammar, and has led to proposals for "verb projection raising" (Haegeman and van Riemsdijk, 1986; Wallenberg, 2009), as well as various clause-union (Haider, 2003, esp. §4), or compound "roll-up" or "remnant" varieties of movement (Koopman, 1996; Koopman and Szabolcsi, 2000), which combinatory derivations like the above reduce to serial contiguous compositional type-dependent constituent merger.

The combinatory derivations shown in figure 21a,b are essentially unaffected by the fact that the arguments are lexically type-raised by case-morphology, which as usual merely reverses the direction of function application mergers. However, the involvement of type-raising allows some

---

[33] As noted in section 2.2, the combination of crossing composition and the second-level composition rules such as (20c) can "grow" categories, and makes the grammar non–context-free. The construction constituted the basis for Shieber's 1985 proof that human languages are not even weakly context-free. In Steedman, 2020:646-7, I examine in greater detail the scrambling possibilities under this generalization for the Germanic constructions exemplified above, following Wurmbrand (2004), and in Hungarian clause union, following Koopman and Szabolcsi.

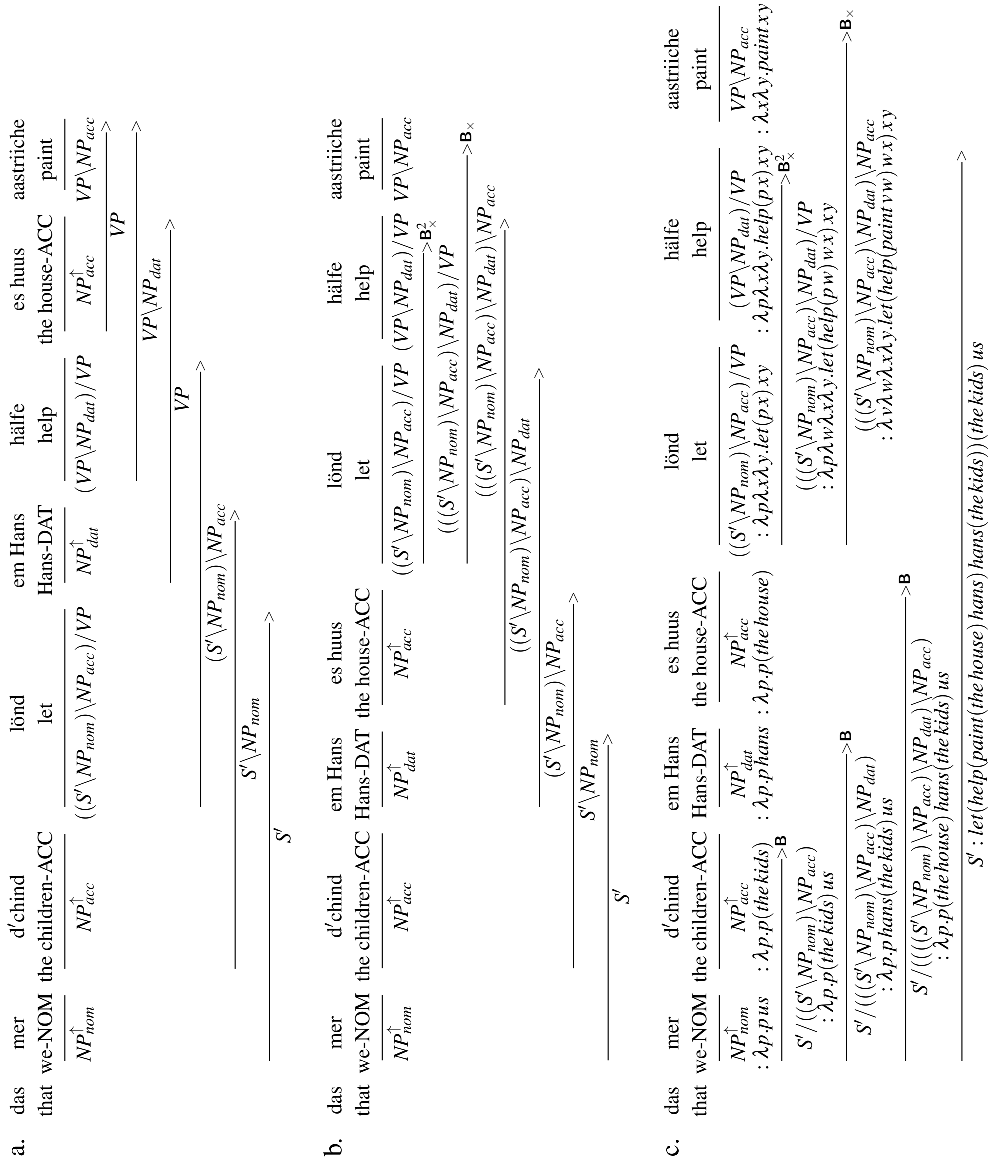

Figure 2: Swiss German: "that we let the children help Hans paint the house" (Shieber, 1985)

further derivations for the latter word-order, because the case-raised argument categories can also compose. The derivation in figure 2c, in which the logical forms omitted in the earlier ones are included, is an example.

The second-order category $S'/((((S'\backslash NP_{nom})\backslash NP_{acc})\backslash NP_{dat})\backslash NP_{acc})$ that is built by successive composition mergers in the latter derivation for the argument-cluster *mer d'chind em Hans es huus*, has the logical form $\lambda p.p\, house\, hans\, children\, us$. This category does the work of Haegeman and van Riemsdijk's and Wallenberg's "verb projection", an argument-structure which Wallenberg derives by multiple head-movements of the verbs to adjoin to their respective vPs, leaving multiple verb traces, and then raises in its entirety to Spec of TP (den Dikken, 1994; Den Dikken, 1995; Wallenberg, 2009:166-167). The present approach differs in making a single second-order bound-variable p in that logical form do the work of their multiple head-movements at the level of lf.

Since the rules of the present grammar are conditioned on syntactic type alone, rather than structure-dependent, they remain entirely blind to the derivation of the composite verb *lönd hälfe aastriiche* of category $(((S'\backslash NP_{nom})\backslash NP_{acc})\backslash NP_{dat})\backslash NP_{acc}$, and to the structure of its lf $\lambda v\lambda w\lambda x\lambda y.let(help(paint\, v\, w)\, w\, x)\, x\, y$, which instantiates the bound variable p in the above argument-cluster lf. Their combination is instead accomplished via a single application merge, rather than via multiple movement, and without the involvement of multiple verb traces (cf. Epstein et al., 1998; Epstein and Seely, 2006:178-179).

The related Dutch bare infinitival verbs *laten*, *helpen*, *zien*, etc., are more restricted, allowing only the crossing dependencies parallel to those in figure 2b,c, and excluding the "canonical" order 2a. Their grammar can be simply captured without rule restrictions, by assuming that they bear a × restriction to crossing composition on their first argument, as in $(VP\backslash NP_{dat})/_{\times}VP$, disallowing the applicative derivation of figure 2a.[34]

The following generalization, of which Wallenberg's condition (45) for Germanic main clauses is an extremely special case, and which can be seen as capturing his more general CoCC, is a prediction of the present theory, rather than a constraint limiting overgeneration.[35]

Let $X\$_1|Z\$_2$ be a (possibly derived) function from zero or more arguments $\$_2$ to a function from Z to a function of zero or more further arguments $\$_1$ into X. Then:

(51) *The Scrambling/Node-raising Generalization:*
An argument $Z^{\uparrow}$ can combine contiguously with $X\$_1|Z\$_2$ to yield $X\$_1\$_2$ iff the order-preserving raised type $X\$_1/(X\$_1\backslash Z)$ or $X\$_1\backslash(X\$_1/Z)$ is among the set of morpholexically licensed type-raised categories $\{Z^{\uparrow}\}$ for the language in question, and there is an applicable combinatory rule among those listed in section 2.2 as licensed by the Combinatory Projection Principle (19).

It is important to notice that the scrambled derivations allowed under this generalization, such as Swiss German figure 2b,c, resemble raising and control, as discussed in section 3, in preserving or "reconstructing" thematic relations at the level of lf predicate-argument structure. Despite wildly nonstandard derivational constituency, the latter reflects more or less traditional notions of constituent structure and command. While Japanese and German main clauses are

---

[34] This result is contrary to the observations of Kuhlmann et al. (2010, 2015), which reflect an earlier version of the theory.

[35] Combinatory scrambling is not permutation-complete for more than three arguments—see Becker et al., 1991; Hockenmaier and Young, 2008; Stanojević and Steedman, 2021 for some discussion.

sometimes described as differing in whether scrambling can "move things out of" finite clauses, this appearance is simply a consequence of exactly which pf alignments can be captured in SOV vs. V2 lexicons.[36]

In summary, morpho-lexical order-preserving type-raising of arguments such as NPs over a language's verbal subcategorization types, interacts productively with function composition (seen in the previous section doing the work of A movement in English), to achieve the effects of scrambling in free- or partially free- word-order languages. The "copies" of the movement theory are again represented in lexical logical forms by binders $\lambda\alpha$ and their variables α, and are projected onto derived forms by rules of strictly contiguous merger, continuing to preserve synchrony between agreement, merger, and semantic transfer. This possibility in turn depends on directionality being specified in the lexicon and projected by rules adhering to the Combinatory Projection Principle (CPP) of section 2.2.

Despite the lexically-imposed restricted opportunities for scrambling in English, its grammar, like that of all languages, is profoundly affected in other ways by the inclusion of merger by composition, and of case, expressed as lexicalized order-preserving type-raising of arguments. In particular, these operations immediately allow an explanation in terms of strictly adjacent "external" merger of constituents (in the extended sense of that term that they engender) for a great diversity of unbounded phenomena that have been attributed to "deletion under coordination" and "*wh*-movement" (Steedman, 1985, *passim*; Phillips, 2003; Bozşahin, 2012), considered in the next two sections.

5 Coordination and Constituency

Conjunctions like *and* carry a category $(\mathrm{X}\backslash_{\star}\mathrm{X})/_{\star}\mathrm{X}$, where X schematizes over S or functors into S with the same low bound on valency as the lexicon (that is, three or possibly four). Its $\star$ slash-types mean that it can only combine by the application rules (11).[37]

In application to the predicate $S\backslash NP$, it allows left-node-raising finite VP coordination, as in figure 3a.

More interestingly, the phenomenon of right-node-raising, including the fact that it is unbounded and can apply to multiple arguments (Abbott, 1976; Gazdar, 1981), is also consistent with the fact that all arguments are order-preserving cased raised types that can compose as well as apply, as we saw in figure 2c for Germanic, as in figure 3b. (Comparison with 3a shows that structurally-cased English needs more than one raised category for objects, as do morphologically-cased languages like Latin, Japanese, and various forms of Germanic.)[38]

As in the discussion of the examples (36) and (38) that were used to introduce type-raising

---

[36] The generalization in (51) is directionally symmetrical, so it is predicted that verb-initial languages/constructions with free word-order should allow mirror-image rightward scrambling. This appears to be the case in some Oceanic languages (Otsuka, 2005). The more widespread nature of leftward displacement in general noted by Kayne (1994) and Cinque (2005), among others, may reflect an information-structural asymmetry between sentence–initial and -final positions, rather than a hard universal constraint against rightward displacement, as noted by Abels and Neeleman (2012).

[37] This category is a necessary consequence of the logical form $\lambda p\lambda q.p\sqcap q$, in which $\sqcap$ is the transitive closure of conjunction $\wedge$ over bounded predicative function types of Partee and Rooth (1983), and p and q are therefore functions of the same bounded semantic type. It is therefore assumed not to be a degree of freedom in the theory that other lexical conjunctions or other languages are free to vary (although they may make finer distinctions, as between nominal or verbal conjunction).

[38] In this and subsequent derivations, the forward and backward applications of the conjunction are abbreviated as a single step, indexed as "$<>$", to save space.

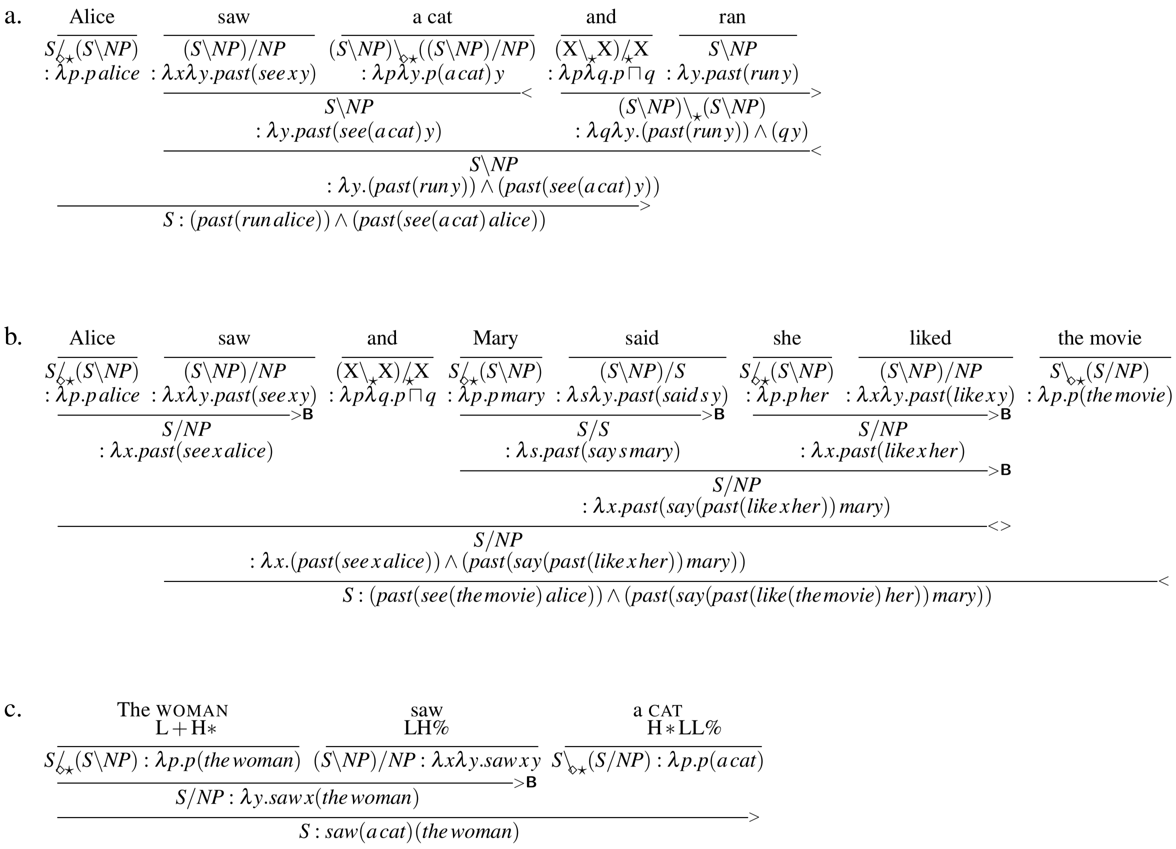

Figure 3: Left- and Right-Node-Raising

in section 4.1, the derivations in figure 3a and b involve only rules of strictly external merger and standard type-raised argument logical forms. However, in passing the values of arguments like *Alice* and *the movie* into the logical form via the second-order variable p, they again have the effect of movement or internal merger, this time as "across-the-board" or "parallel" merge (cf. Citko and Gračanin-Yuksek, 2020 and see section 7 below).

It will also be obvious from derivations like figure 2b,c and 3b that the inclusion of composition and case-raising allows alternative derivations for non-coordinate sentences, with similarly unorthodox constituencies yielding the same logical form as the canonical one. For instance, the alternative derivation for (9) shown in figure 3c is one in which *a woman* first composes with *saw*, and *a cat* subsequently applies to the result, yielding exactly the same lf, $saw(a\,cat)(a\,woman)$.

In the present paper, most derivations will for ease of comprehension be chosen to minimize non-standard constituency of this kind. But such derivations are still available grammatically, as is shown by the possibility of forcing derivations like 3c by applying the indicated "rise-fall-rise" intonational tune L+H* LH% to *The woman saw* to make it into an intonational phrase, say as part of the answer to a question such as *What did the* WOMAN *see?* (Jackendoff, 1972; Pierrehumbert and Hirschberg, 1990; Steedman, 1991a, 2000a,b, 2014b), consistent with a strong form of the MATCH hypothesis of Selkirk (2011) concerning homomorphism of syntax and prosody.

It should be clear from the above example that merger by composition is a free operation of syntax, and is not contingent on movement or displacement of any kind, unlike the notationally similar SLASH feature-passing mechanism of GPSG/HPSG. Under the present theory, *the woman saw* is potentially a constituent of the canonical clause of type S/NP, on the same level as the traditional predicate of type $S\backslash NP$, rather than one of type S bearing a feature indicating the presence of a trace (Gazdar, 1981; Sag, 1997; Jacobson, 1999:149-150, 2014:232).[39]

However, it is important to avoid the error of assuming that non-standard structures like those in figure 3b, and c must allow reflexive binding anomalies, or force wide-scope readings of objects like *every cat* in the analogous derivation for (42). Any phenomenon such as quantifier-scope that reflects c-command must be analysed at the level of lf, where more or less standard relations of command relevant to scope of quantifiers, negation etc., as revealed by bound existentials, NPIs, etc. are, as a consequence of the adjacency assumption (3), preserved by combinatory derivation.[40]

For example, in Steedman 1999, 2000b, 2012, I analyse existentially quantified arguments, as in (42), as generalized Skolem terms, showing how both wide- and narrow-scope readings are predicted for *some movie* in the following RNR example related to figure 3b, as well as the "across the board" exclusion of "mixed" readings with narrow scope in one conjunct and wide scope in the other observed by Geach (1970) (cf. Reinhart, 1997, 2006).[41]

---

[39] Like every other kind of ambiguity, the existence of multiple derivations for the same lf is a problem for the parsing model ("performance"), albeit an entirely solved one—see Eisner, 1996; Hockenmaier, 2003; Clark and Curran, 2007; Lewis and Steedman, 2014. But it is quite irrelevant to our present concern with the theory of grammar ("competence").)

[40] The above error is apparently almost irresistible—see Rothstein (2001); Bernardi (2002, 2004); Uchida (2008); Barker and Shan (2014); Barker (2015), *passim.* See section 6.2 below for implications concerning reconstruction.

[41] See note 26. In Steedman 1996, I discus related issues in the grammar of reflexives and reciprocals, while Steedman (2012) and Jacobson (2014) discuss NPIs and scope of negation as lf phenomena in related categorial frameworks. Other phenomena that have been attributed to surface structural c-command, such as the anti-c-command condition on parasitic gaps (Engdahl, 1983) and intervention effects on *tough*-movement (Keine and Poole, 2017) are shown by Steedman (1987, 1996) to be predicted by CPP. The results of VanLehn (1978); Moulton and Han (2018) and Lidz (2018) showing performance biases of surface order on accessibility of alternate readings are of course irrelevant to the present

(52) Every woman saw and every man said he liked some movie. $\forall\exists/\exists\forall$

Perhaps the most convincing argument for the direct involvement of type-raising in syntax is found in examples like those in figure 4, which have presented a problem for all theories of grammar since the observations of Ross (1967, 1970). They are immediately predicted to be characteristic of English (Dowty, 1985/1988; Steedman, 1985) and Welsh (Borsley et al., 2007:52). (Lf is omitted in Welsh figure 4c, as being entirely parallel to English figure 4, and the derivation of the right conjunct (which is identical to that of the left) is abbreviated): Coordinands like *Mary pizza* and *lyfr i Mair* in figure 4 have exactly the syntactic and semantic character of the raising "verb projection" that we saw for Swiss German in figure 2c. This account of coordination therefore makes strong predictions concerning the interaction of coordination with scrambling in German and Japanese. For example, the non-standard constituents in scrambled derivations like Germanic figure 2c are predicted to "feed" type-dependent coordination, subject to the restriction on valency noted at the beginning of the section, just like all the standard ones in figure 2a, and like the simpler residues of Heavy NP Shift (44) in English:

(53) I [bought from Harry and gave to Mary] a very beautiful book.

In the case of the Swiss-German sentences in figure 2, there are further derivations for the same word-orders which can make any contiguous sub-sequence of verbs and arguments into a constituent that can coordinate subject to the same restriction (Steedman, 1985, 2000b), as predicted by the generalization (51).

The argument and adjunct categories that are involved in cluster coordination of this kind are simply the categories they bear in situ. Ross's 1970 generalization to the effect that rightward arguments and adjuncts universally cluster-coordinate in canonical order to the right, and leftward to the left, is thereby predicted as a consequence of the CPP, 19.

The non-canonical argument sequence in the English Heavy NP Shift construction analysed at (44), as in (54a), is not typable as a constituent, and is not predicted to coordinate:

(54) a. I gave to Harry a book.
b. I gave to Harry a book and to Alice a record.

Nevertheless, (54b) is widely accepted (Beavers and Sag, 2004b), as are the related German scrambled ditransitives. Other cases of non-constituent coordination noted by Beavers and Sag (2004b) are not even limited to sequences of the same type:

(55) a. ?I showed [three boys a video]$_{S/((S/NP)/NP)}$ and [a movie to two girls]$_{S/((S/PP)/NP)}$.
b. ?I showed [a video to three boys]$_{S/((S/PP)/NP)}$ and [two girls a movie]$_{S/((S/NP)/NP)}$

Such examples seem to arise from less strictly type-dependent processes like gapping and stripping that do not seem to be purely syntactic (Hankamer and Sag, 1976; Steedman, 2000b).

Johnson (2017) follows Neijt (1979) in regarding argument/adjunct cluster coordination as arising from a rule of gapping, distinct from left- or right- node-raising. Both authors give as part of their rationale that to assume otherwise would require regarding clusters like *Mary, pizza* as constituents—which of course, under present assumptions is exactly what we should be doing. In Steedman 1990, 2000b, I explain right-conjunct gapping as arising from exactly the same mechanism of composing order-preserving type-raised argument categories as in cluster coordi-

---

concern with competence.

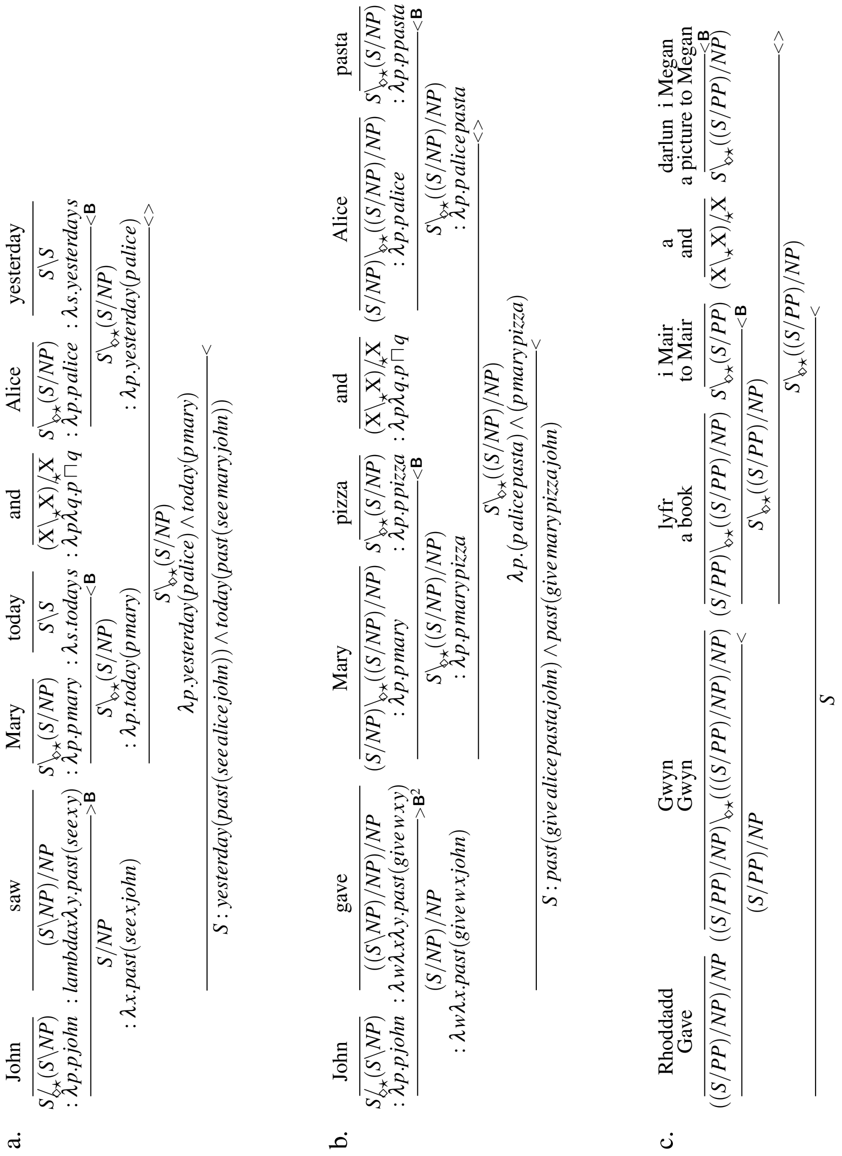

Figure 4: Argument/Adjunct Cluster Coordination

nation, albeit including the subject, and with a separate discourse information-based source for the gap itself.

Hirsch and Wagner (2015) go further than Johnson in taking right-node raising of even a single argument as consequent upon gapping, while Hirsch (2017) seeks to reduce even VP coordination/conjunction reduction, figure 3a, to gapping, also consistent with subject type-raising.

Morpho-lexical type-raising can therefore be thought of as offering a support for these authors' claims concerning the relatedness of these constructions, without endorsing their conclusion that a discontinuous gapping rule per se is their common origin.

Similarly, since the multiple copies of variables that the conjunction category brings to lf in derivations like figure 3a are, as noted earlier, formally equivalent to a directed acyclic graph in which they are represented as a single node, the theory can also be seen as offering a constrained mechanism for establishing multidominance, of the kind assumed by (McCawley, 1982, 1987; Ojeda and Huck, 1987; Phillips, 2003), and Citko (2018; 2017, n.5 notwithstanding).

A number of critics have pointed out that the present theory seems to predict that examples like (56a) below should be grammatical, because the conjuncts *Alice* and *Mary said that John* seem both to be syntactically typable as $S/_{\diamond\star}(S\backslash NP)$:

(56) a. ?[Alice]$_{S/_{\diamond\star}(S\backslash NP)}$ and [Mary said that John]$_{S/_{\diamond\star}(S\backslash NP)}$, liked the movie.
b. ?[Alice said that John]$_{S/_{\diamond\star}(S\backslash NP)}$ and [Mary said that Fred]$_{S/_{\diamond\star}(S\backslash NP)}$, liked the movie.

However, the comparable oddity of (56b) suggests that what is wrong with (56a) is not the types, but some difficulty related to processing associated with forming constituents like *?Mary said that John*, by composition into subjects, such as the fact that English subjects are locally ambiguous as to raised type until the verb is encountered. (Such fragments also seem resistant to being marked as prosodic phrases in the manner exemplified in figure 3c.)

Such a restriction might also explain the "No Embedding" constraint on medial gapping that excludes example like the following (Hankamer, 1979:19—see Johnson, 2017 for discussion):

(57) *[Alfonse]$_{S/_{\diamond\star}(S\backslash NP)}$ stole the emeralds, and [I think that Mugsy]$_{S/_{\diamond\star}(S\backslash NP)}$, the pearls.

As Johnson notes, this constraint does not seem to be universal. Farsi and Hungarian are verb-final languages which allow rightward "SOV and SO" gapping. Farudi, 2013:75ff, Bîlbîie and Faghiri (2022), and Lipták, 2019:835 show that both allow embedded rightward gapping equivalent to (57). Consistent with the above conjecture, both languages are morphologically cased, so that the type of the embedded subject is disambiguated in advance of the verb, unlike English.

The generalization that is captured above is that any contiguous sequence of verbs and/or their arguments and/or adjuncts that can compose to yield a category with bounded valency, including the non-standard constituents involved in *there*-insertion and those discussed in connection with the Scrambling/Node-raising generalization (51) in section 3 and examples like those in figure 2b, are predicted to "feed" the process of coordination of constituents of the same type (cf. Steedman, 1985, 2000b). The combination of composition-merger and morpho-lexical case type-raising thereby allows a wide range of "deletions under coordination" to be analyzed purely in terms of contiguous "external" merger, of the varieties introduced in section 2.2, and as constituent coordination (albeit under an extended definition of constituency), without movement, deletion, extra-lexical copying, syntactic multidominance, "parallel merge", or G/HPSG-style passing of features specific to extraction.

6 *Wh*-extraction

The residues of right-node raising in examples like figure 3b are strikingly similar to the residues of leftward movement in relativization. To analyze unbounded *Wh*-movement as merger, it too must be characterized as application of a second-order function to such a residue. However, the category of *Mary said she liked* in figure 3b is rightward S/NP with respect to its complement NP. The CPP theorems of Consistency (12) and Inheritance (13) exclude syntactic operations that override or alter lexical directionality or its derivational projection onto S/NP, such as FLIP in Williams's (2003) CAT version of categorial grammar. Fronted elements, unlike right-node-raised ones, must therefore bear a non-order preserving raised type of the form $S'/(S|NP)$, where S' is a clausal type distinct from S, such as $S_t, S_{whq}, N\backslash N$, etc.).

As in all cases of leftward and rightward movement, the derivation achieves the effect of internal merger by first composing the residue $S|NP$, which binds an object at lf, using the rules of adjacent external merger listed in section 2.2, then applying the fronted raised type to it in a further external merge to "label" the result S'. By reducing internal merger to external, the present theory avoids anything corresponding to the disjunctive definition of merger in the labeling algorithm of Chomsky, 2008:145. The latter separately stipulates that the internal merger of α with β to form the node or set $\{\alpha, \beta\}$ assigns it the label β, rather than the α stipulated for external merge, a distinction that "strongly suggests a generalization is being missed" (Boeckx, 2015:37).

6.1 Some Simple Cases

Topicalization is the simplest case of move as external merge, where the displaced item carries a category such as $S_t/(S/NP)$, licensed either by sentence-initial position/prosody, as in English, or morphologically, as in Japanese, and marking the resulting S as t(opicalized) (we pass over the information structural semantics of topicalization itself in the lf):

(58)
$$\begin{array}{ccc}
\text{Movies,} & \text{she} & \text{likes}\\
\overline{S_t/(S/NP) : \lambda p.p\,movies} & \overline{S/(S\backslash NP_{3s}) : \lambda p.p\,her} & \overline{(S\backslash NP_{3s})/NP : \lambda x\lambda y.pres(like\,x\,y)}\\
 & \multicolumn{2}{c}{\hrulefill{>}\mathbf{B}}\\
 & \multicolumn{2}{c}{S/NP : \lambda x.pres(like\,x\,her)}\\
\multicolumn{3}{c}{\hrulefill{>}}\\
\multicolumn{3}{c}{S : pres(like\,movies\,her)}
\end{array}$$

Wh-question formation is similar, apart from the fact that *wh*-items other than the subject specify subject-inversion, of which only auxiliaries are heads in English:

(59)
$$\begin{array}{cccc}
\text{What} & \text{does} & \text{she} & \text{like?}\\
\overline{S_{whq}/(S_{inv}/NP)} & \overline{(S_{inv}/VP)/NP_{3s}} & \overline{(S_{inv}/VP)\backslash((S_{inv}/VP))/NP_{3s}} & \overline{VP/NP}\\
: \lambda p\lambda wh.p\,wh & : \lambda y\lambda p.pres(p\,y) & : \lambda p.p\,her & : \lambda x\lambda y.like\,x\,y\\
 & \multicolumn{2}{c}{\hrulefill{<}} & \\
 & \multicolumn{2}{c}{S_{inv}/VP : \lambda p.pres(p\,her)} & \\
 & \multicolumn{3}{c}{\hrulefill{>}\mathbf{B}}\\
 & \multicolumn{3}{c}{S_{inv}/NP : \lambda x.pres(like\,x\,her)}\\
\multicolumn{3}{c}{\hrulefill{>}} & \\
\multicolumn{3}{c}{S_{whq} : \lambda wh.pres(like\,wh\,her)} &
\end{array}$$

(The logical form of a wh-question is simply a λ-term representing a property which is true of entities that answer the question, binding an answer-variable *wh* contributed by the *wh*-item.)

Wh-items such as relative pronouns bear similarly category-changing, non-order preserving type-raised categories. For example:

(60) a. who, that := $(N_{agr}\backslash N_{agr})/(S\backslash NP_{agr}) : \lambda p\lambda n\lambda wh.n\,wh\wedge p\,wh$
b. who(m), that := $(N\backslash N)/(S/NP) : \lambda p\lambda n\lambda wh.n\,wh\wedge p\,wh$

(61)
$$\begin{array}{ccccc}
\text{the} & \text{movie} & \text{that} & \text{she} & \text{likes}\\
\hline
NP^{\uparrow}/N & N & (N\backslash N)/(S/NP) & S/(S\backslash NP_{3s}) & (S\backslash NP_{3s})/NP\\
:\lambda n\lambda p.p(the\,n) & :movie & :\lambda p\lambda n\lambda wh.n\,wh\wedge p\,wh & :\lambda p.p\,her & :\lambda x\lambda y.like\,x\,y\\
 & & & \multicolumn{2}{c}{\overline{\hspace{10em}}{>}\mathbf{B}}\\
 & & & \multicolumn{2}{c}{S/NP:\lambda x.like\,x\,her}\\
 & & \multicolumn{3}{c}{\overline{\hspace{20em}}{>}}\\
 & & \multicolumn{3}{c}{N\backslash N:\lambda n\lambda wh.n\,wh\wedge like\,wh\,her}\\
 & \multicolumn{4}{c}{\overline{\hspace{24em}}{<}}\\
 & \multicolumn{4}{c}{N:\lambda wh.movie\,wh\wedge like\,wh\,her}\\
\multicolumn{5}{c}{\overline{\hspace{28em}}{>}}\\
\multicolumn{5}{c}{NP:\lambda p.p(the(\lambda wh.movie\,wh\wedge like\,wh\,her))}
\end{array}$$

All of these constructions are unbounded, as we saw for related right-node raising (52), because residues like *she says she likes*$_{S/NP}$, *do you regret that you love*$_{S_{inv}/NP}$, and *she said that she likes*$_{S/NP}$, are all derivable with the same type S/NP as *she likes* (see (62) below).

In each case (58-61), the logical form of the *wh*-item "mover" has λ-binder and variable "copies" of p, the "residue of movement" already in place, with the position of the former already specified by its syntactic category $S'/(S|NP)$ as at the left periphery of the latter (where S' is again $S_t, S_{whq}, N\backslash N$, etc.). The binder passes the lf of its argument S/NP to its result as the value of p in an expression such as $p\,apples$ in (58), $\lambda wh.p\,wh$ in (59), and $\lambda n\lambda wh.n\,wh\wedge p\,wh$ in (61), It does not call for *wh*-movers to adjoin to a null C head of CP that is never realized at either of the output levels of phonological or logical form, as in the standard Minimalist analysis criticized by Starke (2001, 2004).

It is in fact the same mechanism that was used in non-movement examples like (42) and figure 3c to bind in situ arguments to their theta-positions. This interpretation of copying is a therefore a theory of selection and merger in general, rather than of movement alone.

6.2 "Pied-piping" and "Late Merge" as Multiple Reconstruction

More complex "pied-piping" *wh*-items like *which movie* or "the height of the lettering on the covers of which" inherit this category from *wh* words, just as ordinary nominals do from determiners like (40) and (41) (Steedman, 1996, 2012:89-91). For example:

(62)
$$\begin{array}{ccccccc}
\text{Which} & \text{movie} & \text{does} & \text{she} & \text{say} & \text{she} & \text{liked?}\\
\hline
(S_{whq}/(S_{inv}/NP))/N & N & (S_{inv}/VP)/NP_{3s} & NP^{\uparrow}_{3sf} & VP/S & NP^{\uparrow}_{3sf} & (S\backslash NP)/NP\\
:\lambda n\lambda p\lambda wh.n\,wh\wedge p\,wh & :movie & :\lambda p\lambda y.pres(p\,y) & :\lambda p.p\,her & :\lambda s\lambda y.say\,s\,y & :\lambda p.p\,her & :\lambda x\lambda y.past(like\,x\,y)\\
\multicolumn{2}{c}{\overline{\hspace{12em}}{>}} & \multicolumn{2}{c}{\overline{\hspace{12em}}{<}} & & \multicolumn{2}{c}{\overline{\hspace{12em}}{>}\mathbf{B}}\\
\multicolumn{2}{c}{S_{whq}/(S_{inv}/NP)} & \multicolumn{2}{c}{S_{inv}/NP} & & \multicolumn{2}{c}{S/NP}\\
\multicolumn{2}{c}{:\lambda p\lambda wh.movie\,wh\wedge p\,wh} & \multicolumn{2}{c}{:\lambda p\lambda x.pres(p\,her)} & & \multicolumn{2}{c}{:\lambda x.past(like\,x\,her)}\\
 & & \multicolumn{3}{c}{\overline{\hspace{16em}}{>}\mathbf{B}} & & \\
 & & \multicolumn{3}{c}{S_{inv}/S:\lambda s.pres(say\,s\,her)} & & \\
 & & \multicolumn{5}{c}{\overline{\hspace{30em}}{>}\mathbf{B}}\\
 & & \multicolumn{5}{c}{S_{inv}/NP:\lambda x.pres(say(past(like\,x\,her))\,her)}\\
\multicolumn{7}{c}{\overline{\hspace{44em}}{>}}\\
\multicolumn{7}{c}{S_{whq}:\lambda wh.movie\,wh\wedge pres(say(past(like\,wh\,her))\,her)}
\end{array}$$

The logical form that results from derivation (62) results from applying the wh-item lf $\lambda p\lambda wh.movie\,wh\wedge p\,wh$, to the potentially unbounded lf-residue of relativization corresponding *does she say she liked*—that is, $\lambda x.pres(say(past(like\,x\,her))\,her)$. The result is to "reconstruct" the λ-bound answer-variable *wh* as the object of *like*, c-commanded by both its subject and that of *say*, namely *her*.

However, reconstruction in this sense does not involve the entire *wh*-item, as seems to be standardly assumed under the copy theory. In particular, the pied-piped noun property *movie* is not carried along with the answer variable *wh*, which is separately "reconstructed" as its argument, as under relativization in (61).

Von Stechow (1996) argues for a related derivational analysis, in which covert movement of *wh* out of the overtly-moved *whose books* is achieved at "an intermediate level between S-structure and LF . . . called WH-structure . . . followed by reconstruction at LF." (Von Stechow's covert movement analysis is endorsed by Sauerland and Heck (2003) and Lechner (2013):169, who discuss a number of LF-Intervention effects in support.) In present terms, von Stechow's WH-structure and the domain of covert movement are both latent in the lexical logical form of *wh*-words, and the dependencies defined there by λ-bound *wh*.

The present analysis resolves a notorious difficulty for the binding theory. Standard reconstruction of the full moved item "which book that Harry reviewed" (or its lf) to object position in examples like the following would lead to a Condition C violation:

(63) Which book that Harry$_i$ reviewed did he$_i$ like ~~*which book that Harry$_i$ reviewed*~~*?*

This observation has led to a plethora of proposals for "optional" reconstruction, free type-lifting or "continuation-passing", "Roll-up" movement, "Late Merge" of adjuncts, or definitions of binding in terms of "Almost C-command", "Partially-determined Full Interpretation", "bleeding" Condition C, higher-type $((e \to t) \to t)$ traces, and/or "hybrid" movement systems involving both copies and traces (Reinhart, 1983; Higginbotham, 1983; Lebeaux, 1991, 2009; Jacobson, 1994; Chomsky, 1995b:200-204; Hornstein, 1995:108-111; Brody, 1995:129-144; Romero, 1998; Fox, 1999, Takahashi and Hulsey, 2009; Barker, 2012; Bruening, 2014; Lechner, 1998, 2018; Keine and Poole, 2018).

According to the theory proposed here, by contrast, the logical form of (63) is the following (the derivation is suggested as an exercise):

(64) $\lambda wh.((book\, wh) \wedge (review\, wh\, harry)) \wedge (like\, wh\, him)$

While the lf "reconstructs" the λ-bound answer-variable *wh* into argument position in the individual predications that mean the answer *wh* is a book, that Harry reviewed it, and he liked it, there is no sense in which *him* c-commands its antecedent *harry* at lf.[42]

Of course, neither does *harry* command *him* at lf in (64), so the way pronouns actually do access their referents remains to be explained. However, examples like (63) suggest that the relation between *Harry* and *he* is not determined by structural command or variable binding at any level, and that Condition C and its "obviation" or "bleeding" under reconstruction reflects possibilities for coreference to a dynamic discourse-model that have nothing directly to do with c-command or true binding of the kind seen with quantifier-bound variables, of a kind proposed by Postal (1966), Wasow (1972), Stalnaker (1974), Lewis (1979b), Kamp (1981/1984), and Kuno (1987), and discussed by Chierchia, Jacobson, and Bruening. Such a conclusion seems consistent with the referentiality effects on reconstruction possibilities noted by Reinhart and Reuland (1991) and Heycock (1995), and with the resemblance of the lf conjuncts in (64) to

---

[42] On the assumption that related examples involving complements like "Which claim that Harry$_i$ is a genius does he$_i$ endorse?" have lfs like the following, similar avoidance of Condition C violations is predicted for them, contra Chomsky, 1995b:204 and Brody, 1995:140:

(i) $\lambda wh.claim\,(genius\, harry)\, wh \wedge endorse\, wh\, him$

"phases" (Chomsky, 2001). Clearly, there is more to say about this open problem than there is room for here (cf. Büring, 2005:256-260; Bruening and Al Khalaf, 2019; Sportiche, 2019).

6.3 Islands

The various types of long-range dependency considered in the preceding sections have one characteristic in common: In every case, the element we think of as displaced is an argument, such as *NP*, albeit with a raised type such as $S_t/(S/NP)$, and the residue of displacement is a function over that type, which must be derived by iterated function composition.[43]

If a category is an adjunct, such as the adverbial *whistling "Dixie"*, $VP\backslash VP$, then the fact that adjuncts select for the VP, rather than being selected by it as an argument, means that no rule permitted by the Combinatory Projection Principle of Consistency (12) of section 2.2 will allow a category of the form $(\mathrm{X}\backslash\mathrm{X})/\mathrm{Y}$ (such as the adverbial head *whistling* $(VP\backslash VP)/NP$) to combine with its second argument X (such as the VP *work*) to the left in advance of combination with Y (such as the *NP "Dixie"*) to the right.

Under the present definition of merger, adjuncts are therefore predicted to be islands to extraction, as observed by Huang, 1982:505, following Cattell (1976), and many others:

(65) a.
$$\begin{array}{ccccc}
\ast\text{What tune} & \text{does} & \text{Harry} & \text{work} & \text{whistling?} \\
\hline
S_{whq}/(S_{inv}/NP) & (S_{inv}/VP)/NP_{3s} & NP^{\uparrow}_{3s} & VP & (VP\backslash VP)/NP \\
\cline{2-3}\cline{4-5}
 & \multicolumn{2}{c}{S/VP \quad <} & \multicolumn{2}{c}{\ast\ast\ast \quad ***}
\end{array}$$

b. *Who did Harry [file the reports]$_{VP}$ [without telling]$_{(VP\backslash VP)/NP}$?

c. *Who do you know a [man]$_{N}$ [that met]$_{(N\backslash N)/NP}$?

Because all arguments are assumed to be lexically type-raised, in most cases to the exclusion of the corresponding unraised category, and type-raised categories $X^{\uparrow}$ are VP-adjunct-like, in the sense that they select verbs and verb-phrases as in $VP|(VP|X)$ etc, we correctly predict that raised arguments including $NP^{\uparrow}$ are also islands:[44]

(66) *What do you [doubt]$_{VP/NP}$ [the claim that he has read]$_{(VP\backslash(VP/NP))/NP}$?

The possibility of explaining island effects in this way depends crucially on the fact that linear correspondence is specified in the lexicon, and that type-raising is lexical rather than syntactic.

In English, S, $\overline{\mathrm{S}}$, and the various kinds of VP complements are the only phrasal arguments that exist in their unraised form. However, their participation in argument-cluster coordinations

---

[43] This generalization holds for "remnant" movement, such as Germanic transitive verb-topicalization, in which the OV verb is a (crucially, rightward) argument of the type-raised accusative:

(i)
$$\begin{array}{cccc}
\text{Essen} & \text{wird} & \text{er} & \text{Äpfel} \\
\text{Eat} & \text{will} & \text{he} & \text{apples} \\
\hline
S_t/(S_{inv}/(VP\backslash NP)) & (S_{inv}/VP)/NP_{3s} & (S_{inv}/VP)/((S_{inv}/VP)/NP_{3s}) & VP/(VP\backslash NP) \\
\cline{2-3}
 & \multicolumn{2}{c}{S_{inv}/VP \quad <} & \\
\cline{2-4}
 & \multicolumn{3}{c}{S_{inv}/(VP\backslash NP) \quad >\mathbf{B}} \\
\cline{1-3}
\multicolumn{3}{c}{S_t \quad >} &
\end{array}$$

[44] Certain constructions that look like extractions from NP and PP like the following seem to arise from lexicalized multi-word expressions or verb-particle constructions:

(i) a. Who did they take advantage of?

b. What will you paint a picture of?

The relative weakness of "derived" subject islands such as unaccusatives found by Jurka (2013) and Polinsky et al. (2013) remains unexplained in these terms.

like the following, analogous to those in figure 3, suggests that they also bear adjunct or raised categories:

(67) I will tell$_{(VP/S)/NP}$ [[[Donald] [(that) he is fired]] and [[Ronald] [(that) he is hired]]]$_{VP\backslash((VP/S)/NP)}$.

The categorial ambiguity claimed here for English complements is clearly a lexical degree of freedom upon which languages can be expected to differ, some including sentential complements that bear only adjunct or type-raised categories, making them islands, as appears to be the case for *daß*-complements in many Northern dialects of German, or including different complement types, some of which are adjunct-only, and others subcategorized-for arguments, as is the case for *wh/that*-complements in English:

(68) a. *Who were you surprised when you saw?
b. Who were you surprised that you saw?

If a language like English can arrange its lexicon so as to make certain components such as *that*-complements bear both adjunct/type-raised (island) and complement (non-island) categories, it is clear that we must expect islands in general to appear to exhibit a continuum of extractability, from "strong" islands bearing only adjunct categories that are not subcategorized-for and completely block extraction, to "weak" islands bearing both adjunct and argument categories, the latter sometimes subcategorized for by particular verbs. Those verbs and no others can therefore compose into the argument category, allowing extraction (Cinque, 1990; Szabolcsi, 2007; Truswell, 2007b,a; Boeckx, 2012:16). Truswell illustrates the strong/weak distinction in minimal pairs like the following, among many others:

(69) a. *What tune does John work whistling?
b. What tune did John drive Mary crazy whistling?

Example (69a), repeated from (65a), shows that VP-modifiers like *whistling "Dixie"* are not subcategorized-for by predicates like *work VP*, with which they can only combine as adjuncts $VP\backslash VP$, which block extraction as in (65). However, (69b), shows that *whistling "Dixie"* also carries the argument category VP_{ing}, allowing extraction past subcategorizing verbs, as in *What tune is John whistling?*. In the case of (69b), this implies a category $((VP/VP_{ing})/AP)/NP$ for *drive* and related causatives like *make*, subcategorizing for VP_{ing}, and thereby allowing (69b) by composition of *drive Mary crazy*$_{VP/VP_{ing}}$ into the other category for *whistling*, VP_{ing}/NP.

These observations mean that when we talk of modifiers like *whistling "Dixie"* as "weak islands", we simply mean that they are lexically ambiguous. They are strong islands under their adjunct category, and non-islands under their argument category—but only for matrix verbs that actually specify that category.

This means of course that *John drives Mary crazy whistling "Dixie"* is ambiguous between a (preferred) argument reading in which it is specifically John's whistling "Dixie" that drives Mary crazy, and an adjunct reading analogous to *John works whistling "Dixie"*, where John merely happens to whistle that tune while doing whatever it is that actually drives her crazy.

The exact conditions under which weak island ambiguities are resolved in favor of the complement category, or novel complement-specifying verbs are accommodated via the usual process of adult lexical acquisition, to permit such extractions, remain unclear. They depend upon the matrix-verb's subcategorization(s), the parsing model, and/or world knowledge, as proposed in "connectionist" neural-computational terms by Dowty, 2003:60 and in event-semantic terms by Szabolcsi and Zwarts (1993) and Truswell, rather than upon syntax alone.

6.4 Celtic Complementation: Unbounded vs. Cyclic Movement

Scots Gaelic, like other Celtic languages, has a relative clause marker *a*. However, all authorities insist that Gaelic, Irish, and Welsh *a* is not a relative pronoun, parallel to English *wh*, but a complementizer (McCloskey, 1979; Gillies, 1993; Borsley et al., 2007).

Gaelic also has neutral complementizer *gu(n)*. The two complementizers are in complementary distribution: *gu(n)* never acts as a relative marker, and *a* cannot function as a sentential complementizer for verbs like *abair* ("say").

When *a*-relativization is long-range, the embedded complementizer(s) must be *a*, not *gu(n)*. That is, *gu(n)* complements are islands, parallel to Northern German *daß*, and to Irish *go*: Adger notes the following pattern:

(70) a. $\ldots gu(n) \ldots gu(n) \ldots$ (complementation)
b. $*\ldots gu(n) \ldots a \ldots$ (*)
c. $\ldots a \ldots a \ldots t$ (*wh*-relativization)
d. $*\ldots a \ldots gu(n) \ldots t$ (*)

McCloskey (1990, 2017), Adger (2003); Adger and Ramchand (2005), and Boeckx (2008) have claimed that this pattern of relative clause formation, which also appears in Irish, shows that movement is necessarily cyclic and hence derivation-dependent. because unbounded movement would not be able to "notice" the presence of an island barrier *gu(n)* in pattern (70d).

However, under the present account, the residue of movement must be composed by local mergers, including those involving complementizers. The fact that *gu(n)* is a barrier simply means that *gu(n)*-complements are type-raised as $S\backslash(S/S')$, and hence islands, like *daß*-complements in many Northern German dialects. By contrast, *a* is an unraised complementizer S_a/S, which can be composed into like English *that*.

We pass over the further details here for reasons of space. Like any island effect, movement is blocked because the residue of relativization cannot form in the first place. The mover *a* has no need to "notice" why not.[45]

7 Symmetry and Asymmetry in Left- and Right- Extraction

The reduction of both *wh*-extraction and right-node–raising to contiguous leftward or rightward adjacent merge depends on the possibility of making the residue of both into a constituent of type $S|X$ by identical processes of function composition. The present theory, like some other Minimalist accounts and the earliest version of GPSG (Gazdar, 1981), therefore makes a broad prediction of symmetry for left- and right-extraction in English: whatever can undergo *wh*-extraction from the periphery of a typable constituent can potentially right-node–raise from it, and vice versa, broadly subject to the same constraints, as Boeckx 2015:38 points out.

For example, both leftward and rightward extraction are predicted to be subject to the Coordinate Structure Constraint (CSC), the "Across-the-Board" (ATB) exception to CSC, and the

---

[45] The pattern in Irish is similar to Gaelic, but is complicated by the possibility of resumptive relativization. A related analysis appears to be applicable to the Germanic "*wh*-copying" phenomenon discussed by Felser (2004).

Same-Case Condition on the ATB exception to CSC (Ross, 1967; Williams, 1978; Gazdar, 1981; Pesetsky, 1982; Sag et al., 1985; Steedman, 1985, 2000b), according to the following pattern:[46]

(71) a. Which movie did she say [she liked and Harry disliked]?
b. *Which movie did she say[she liked and Harry disliked the play]?
c. *Which movie did she say[she liked the play and Harry disliked]?
d. *Which movie did she say[she liked and annoyed Harry]?

(72) a. [She said she liked and Harry said he disliked] the movie.
b. *[She said she liked and Harry said he disliked the play] the movie.
c. *[She said she liked the play and Harry said he disliked] the movie.
d. *[She said she liked and annoyed Harry] the movie.

ATB extraction has been attributed by Nunes (2004:127-130) and Zhang (2010) to sideward movement of *which movie* from the right conjunct to the left, although Nunes notes that this does not explain the CSC itself and the unacceptability of (71b,c, and d). Like the related phenomena of parasitic *wh*-movement and adjunct control (25), which Nunes also attributes to sideward movement, the behavior illustrated in (71) follows instead, as in GPSG, from the fact that the category $(\mathrm{X}\backslash_{\star}\mathrm{X})/_{\star}\mathrm{X}$ of conjunctions limits them to combining categories of like type. As for Gazdar (1981), both the CSC and the sideward-movement/parallel-merge case-dependent ATB exception to it are therefore corollaries of the lexicalization of coordination as type-compatibility, rather than arising from structure-dependence in rules or extrinsic parallelism constraints, as proposed in Goodall, 1987; Zhang, 2010; Citko, 2012, and Citko and Gračanin-Yuksek, 2020.

Nevertheless, despite the general tendency to symmetry between left- and right-extraction possibilities, there is an important difference between them. Whereas rightward movers bear the same order-preserving type-raised category as in-situ arguments, English left-extracting categories such as the relative pronoun $(N\backslash N)/(S|NP)$ differ in being non-order-preserving This difference permits certain exceptions to symmetry that we turn to next, all of which must under present assumptions be specified in the English lexicon, and, as such, are expected to vary across other languages.

7.1 Embedded Subject Extraction

One very compelling left-right asymmetry arises from the fact that English embedded subjects can *wh*-extract from the bare complements selected by a very restricted class of verbs like *think*, while the corresponding rightward movement remains impossible:

(73) a. Which critic did you say likes the movie?
b. *Harry thinks walks and Alice says talks, a Boojum.

The possibility of (73a) arises from the possibility of crossing composition of the verb into the complement:[47]

---

[46] Some apparent exceptions to ATB *wh*-extraction like the following were noted by Ross (1967); Goldsmith (1985); Lakoff (1986); Munn (1999); Kehler (2002), and Asudeh et al. (2002):
(i) This is the stuff that those guys in the Caucasus drink all day, and live to be a hundred.
They are discussed by Postal (1998:Ch.4) and by Steedman (2012:94-95), who point out that the examples in question are heavily pragmatically loaded.

[47] This analysis of subject extraction is different from the one presented in earlier publications since Steedman (1987). I am indebted to Haixia Man for discussions of subject extraction in Chinese that led me to it.

(74)
$$\begin{array}{cccc}
\text{Which critic} & \text{did you} & \text{say} & \text{likes the movie?}\\
\overline{S_{whq}/(S_{inv}\backslash NP_{3s})} & \overline{(S_{inv}/VP} & \overline{VP/S} & \overline{S\backslash NP_{3s}}\\
 & & \multicolumn{2}{c}{\overline{\hspace{4cm}}{>}\mathbf{B}_{\times}}\\
 & & \multicolumn{2}{c}{VP\backslash NP_{3s}}\\
 & \multicolumn{3}{c}{\overline{\hspace{6cm}}{>}\mathbf{B}_{\times}}\\
 & \multicolumn{3}{c}{S_{inv}\backslash NP_{3s}}\\
\multicolumn{4}{c}{\overline{\hspace{10cm}}{>}}\\
\multicolumn{4}{c}{S_{whq}}
\end{array}$$

Since the residue of embedded subject relativization bears the same category as that of root subject relativization, the possibility of ATB "left-node-raising" multiple subject extractions like the following is predicted:

(75) A critic that [[panned the movie,]$_{S\backslash NP}$ but [you said liked it.]$_{S\backslash NP}$]$_{S\backslash NP}$

However, the residue of embedded subject *wh*-movement bears the leftward-looking category $S\backslash NP$. Right node raising arguments simply bear the standard order-preserving category of rightward arguments, so RNR of embedded SVO subjects (73b) remains impossible.

Potential overgeneralizations like the following are ruled out for the same reason as (50) in German and the possibility of verb-fronting in English (see note 43): there is no lexical verb-type, and therefore no morpholexically cased type-raised NP, that would allow the derivations:

(76)
$$\begin{array}{cccc}
{*}\text{Did you} & \text{Harry} & \text{say} & \text{likes the movie?}\\
\overline{S_{inv}/VP} & \overline{VP\backslash(VP/NP)} & \overline{VP/S} & \overline{S\backslash NP_{3s}}\\
 & & \multicolumn{2}{c}{\overline{\hspace{4cm}}{>}\mathbf{B}_{\times}}\\
 & & \multicolumn{2}{c}{VP\backslash NP_{3s}}\\
 & \multicolumn{3}{c}{\overline{\hspace{7cm}}{***}}\\
 & \multicolumn{3}{c}{***}
\end{array}$$

The Fixed Subject Effect (Bresnan, 1972) or "that-*trace*" Filter (Chomsky and Lasnik, 1977) can then be captured lexically by limiting the complementizers in question using $/_{\diamond\star}$ slash-type to disallow crossing composition without appeal to any constraints such as the Empty Category Principle (ECP) of Chomsky (1981), the Generalized Left-Branch Condition (GLBC) and related SLASH Termination Metarules of GPSG (Gazdar, 1981:161, Gazdar et al. (1985):161), or the Trace Condition of HPSG (Pollard and Sag, 1994:172-3):

(77)
$$\begin{array}{cccc}
{*}\text{Which critic} & \text{did you say} & \text{that} & \text{likes the movie?}\\
\overline{S_{whq}/(S_{inv}\backslash NP)} & \overline{S_{inv}/S} & \overline{S/_{\diamond\star}S} & \overline{S\backslash NP_{3s}}\\
 & \multicolumn{3}{c}{\overline{\hspace{7cm}}{***}}\\
 & \multicolumn{3}{c}{***}
\end{array}$$

Nevertheless, if sentential adverbials like *in your opinion* are less restricted S/S, subject extractions like the following, noticed by Culicover (1993), are correctly predicted:

(78)
$$\begin{array}{ccccc}
\text{Which critic} & \text{did you say} & \text{that} & \text{in your opinion} & \text{likes the movie?}\\
\overline{S_{whq}/(S_{inv}\backslash NP_{3s})} & \overline{S_{inv}/S} & \overline{S/_{\diamond\star}S} & \overline{S/S} & \overline{S\backslash NP_{3s}}\\
 & & \multicolumn{2}{c}{\overline{\hspace{3cm}}{>}\mathbf{B}} & \\
 & & \multicolumn{2}{c}{S/S} & \\
 & & \multicolumn{3}{c}{\overline{\hspace{6cm}}{>}\mathbf{B}_{\times}}\\
 & & \multicolumn{3}{c}{S\backslash NP_{3s}}\\
 & \multicolumn{4}{c}{\overline{\hspace{8cm}}{>}\mathbf{B}_{\times}}\\
 & \multicolumn{4}{c}{S_{inv}\backslash NP_{3s}}\\
\multicolumn{4}{c}{\overline{\hspace{9cm}}{>}} & \\
\multicolumn{4}{c}{S_{whq}} &
\end{array}$$

The above account correctly predicts that embedded subject-object extraction asymmetries are not characteristic of verb-final and verb-initial languages (Chung, 1983; Maling and Zaenen, 1978): either both will extract, or neither will, as in the various German dialects.

7.2 Right-edge Restrictions

The following apparent asymmetry in English dative extraction led Wilder, 1999: ex.(5) to postulate a Right Edge Restriction (RER) specifically on right-node–raising (79a), from which leftward *wh*-extraction (b) is exempt (the judgments are Wilder's):[48]

(79) a. *I gave a present and congratulated all the winners.
b. The man who I gave a present and congratulated.

Under the current theory, according to which the residues of *wh*-constructions and RNR are identically formed, either both are predicted to to be in, or both out. Under the theory presented so far, the latter is the case, for reasons discussed in connection with example (43).

This prediction is consistent with early transformational accounts of the English ditransitive by North American linguists, who assumed that even the simplest *wh*-extraction of datives was disallowed (Fillmore, 1965; Oehrle, 1976), a restriction which would exclude both (79a) and (b). The fact that many speakers nevertheless accept dative relatives seems to suggest the availability for such speakers of an alternative, possibly deprecated, scrambling lexical category for ditransitives, allowing derivations like the following, and allowing both (79b) and (a), as well as *?I gave a present all the winners*, otherwise excluded as in (43):

(80) ?The girl

$$\begin{array}{ccccc}
\text{that} & \text{I} & \text{gave} & \text{flowers} & \\
\hline
(N\backslash N)/(S/NP) & S/(S\backslash NP) & ?((S\backslash NP)/NP)/NP & ?(S\backslash NP)\backslash((S\backslash NP)/NP)) & \\
\lambda p\lambda n\lambda x.nx\wedge px. & \lambda p.pme & \lambda x\lambda w\lambda y.gave\,wxy & \lambda p.p flowers & \\
\cline{3-4}
 & & \multicolumn{2}{c}{?(S\backslash NP)/NP : \lambda x\lambda y.gave flowers\,xy} & <\mathbf{B}_{\times} \\
\cline{2-4}
 & \multicolumn{3}{c}{?S/NP : \lambda x.gave flowers\,xme} & >\mathbf{B} \\
\cline{1-3}
\multicolumn{3}{c}{?N\backslash N : \lambda n\lambda x.nx\wedge gave flowers\,xme} & & >
\end{array}$$

While this analysis allows the to my ear marginal (79a), it captures the also somewhat marginal (79b), while continuing to allow Heavy NP shift examples like the following, which any more general restriction on non-peripheral rightward movement would otherwise appear to be in danger of excluding:

(81) a. I saw yesterday and congratulated all the winners.
b. I sold to the library and Mary donated to the museum several very valuable books.

To the extent that Wilder's asymmetry (79) is real, it may reflect a differing degree of semantic compatibility of full-arguments and *wh*-elements when called on to simultaneously satisfy dative and accusative rules across the two conjuncts, rather than extractability as such.

7.3 Asymmetric Islands

There have been many claims in the literature since Wexler and Culicover (1980:299) that left- and right- movement are similarly asymmetrical with respect to island constraints—but in the opposite direction to subject extraction and RER above, with RNR the more permissive (Beavers and Sag, 2004a; Cann et al., 2005; Sabbagh, 2007; Citko and Gračanin-Yuksek, 2020: but see

---

[48] Related examples were the reason for GPSG abandoning Gazdar's 1981 claim that RNR was mediated by the same mechanism as *wh*-extraction, and led to HPSG's uneasy embrace of a deletion/ellipsis analysis (Beavers and Sag, 2004a).

Postal, 1998, Steedman, 2012:101-103, Bachrach and Katzir, 2009, and Kubota and Levine, 2020:325-327 for counter-arguments.) Examples like the following are not entirely compelling (the judgments are Beavers and Sag's), especially when care is taken to make intonational prosody the same in (a) as in (b) (Steedman, 2000a, 2012:103):

(82) a. ??Those unflattering pictures of Qaddafi, Yo knows several men who buy, and Jan knows several men who sell.
b. Yo knows several men who buy, and Jan knows several men who sell, those unflattering pictures of Qaddafi.

It seems possible that the asymmetries between *wh*- and RNR–extraction claimed (with some degree of uncertainty) by Citko and Gračanin-Yuksek (2020:ch.3) for Slavic languages are similarly discourse-sensitive, rather than reflecting any fundamental difference in the nature of the long-range dependencies involved. Sabbagh, citing an anonymous reviewer, offers the following apparent RNR exception to the adjunct island constraint:

(83) Politicians win when they defend, and lose when they attack, the right of a woman to an abortion.

Again, the judgment is his, but the corresponding left-extraction *What right do politicians win when they defend and lose when they attack?* seems no worse, especially when the same prosody is applied. The same seems to be true of the "non-coordinate RNR" examples discussed by Hudson (1976); Postal (1994), and Phillips (2003). Bachrach and Katzir (2009) and Hirsch and Wagner (2015) discuss related examples of ATB *wh*-extraction out of islands. The lack noted earlier of a clear distinction between strong and weak islands, and the dependence of the latter on pragmatic factors, make it hard to draw any firm conclusion from these data.[49]

8 Conclusion: On Locality

The notion of locality in syntactic relations that is proposed in this paper is defined lexically, as the domain of a binarized head, such as a verb, specifying a number of co-arguments, with no distinction between the extracted or in-situ status of the latter. All of the syntactic combinatory rules listed in section 2.2 are by definition strictly local, by virtue of the adjacency assumption (3) and the string-adjacency–based definition of the slash directionality feature in section 2.1. Those rules are thereby conditioned only on the syntactic categorial types of string-adjacent contiguous constituents, rather than on their derivational structure or logical form. They need no constraints on their operation other than those projected from lexical types.

Surface syntactic discontinuity in all its forms can thereby be analyzed in terms of strictly local, contiguous merger of local domains. Its appearance in natural languages arises in every case from the merger of a second-order functor, (such as a structurally- or morphologically-

---

[49] Beavers and Sag, 2004a also note, following Davis (1992), that in Hausa, an SVO language with object pro-drop, while ATB object *wh*-extraction is allowed, RNR is not (Davis, 15, 16). However, the availability of object pro-drop in Hausa means that phonological emptiness cannot be taken as evidence of movement per se. If Hausa "movement" is really left-dislocation with pro-drop then there may be asymmetries with respect to discourse characteristics of left- and right-dislocated elements, with the former being by definition discourse-old as required by pro, but the latter required to be new or contrastive, as in English RNR. (This suggestion seems consistent with Davis's own analysis of finite and non-finite verbs in Hausa (ibid:ex.(22)).) Related considerations may explain the asymmetry noted by McCloskey (1986) for Irish prepositions, which engender obligatory pro-drop, and "strand" for right-node raising, but not for *wh*-extraction (although *pro* can act as resumptive—see Legate 1999:ex.(11)).

cased argument, *wh*-item, conjunction, raising verb and/or existential *there*), with its argument, a contiguous first-order functor defining the domain of dependency projection.

That domain is originally established via λ-binding in lexical logical forms. Such bindings are projected onto derived categories by a succession of contiguous mergers, possibly including composition across tensed clause-boundaries, to create potentially unbounded dependency domains, without requiring cyclic mediation via any COMP "escape-hatch". The possibility of such mergers is in general entirely independent of whether extraction is involved or not.

"Internal" Merger is thereby reduced at the level of syntactic derivation to adjacent or "External" Merger, delivering on the promise of Epstein et al. (1998), Chomsky (2001/2004, 2007, 2019), and Epstein and Seely (2006) to reduce MOVE to purely local contiguous merger.

Such mergers build derived categories monotonically, projecting unchanged the syntactic and semantic dependencies originally established in the lexicon, in conformity with the Inclusiveness and Extension Conditions. In contrast to related Minimalist systems such as Jackendoff (1997) and Chomsky (2001/2004), the counterparts of AGREE, MERGE, and phonological and semantic composition or TRANSFER, are entirely synchronous. As well as dynamic movement, the related apparatus of "probes", "goals", and "valuations" can be eliminated, together with attendant "feature deletion" and "visibility" conditions (Radford, 2004:289). Both the "external" and "internal" clauses of the labeling algorithm of Chomsky, 2008:145, are deterministic consequences of the combinatory rules and the (first- or second–order) lexical types that they project, without any possibility arising for "deviant" labeling. Different varieties of movement, such as roll-up/remnant, sideward, parallel, and head varieties, are are all reduced to contiguous merger.

Discontinuity is "translated" by λ-bindings in the same sense as Heim and Kratzer's *Traces* rule (1998:97), Fox's λ-calculus-based *Trace Conversion* (2002:67), and Adger and Ramchand's abstraction mechanism (2005:170-173). The difference is that these dependencies originate as local λ-bindings in lexical logical forms and are directly projected as a consequence of the surface-compositional derivation itself (cf. Epstein and Seely, 2006:7-8,178-180). While a variable α and its binder $\lambda\alpha$ do the work of copies in identifying the source and target of long-range dependency, they do so simply as a projection of the mechanism that binds local in situ complements to lexical heads in derivations like (9) and (42).

As a consequence, derivation structure can be entirely eliminated as a representational level, along with attendant processes of translation to logical form. The sole structural level of representation is lf, but it is a representation to which syntactic derivation itself is entirely blind. (The results of "covert" movement, such as QR (including "long" QR) and antecedence, are already established at this level, having been either projected, like wh-bindings, from lexical determiners like (41) (Montague, 1973; Steedman, 2012), or established dynamically (Chierchia, 1995; Jacobson, 2014)).[50]

As noted earlier (see n.17), the combinatory rules of merger are of low "near-context-free" expressive power. The problem of language acquisition is thereby simplified. The Categorial and Adjacency assumptions (2) and (3) of section 2 and their corollary the CPP (19) generate all and only the universal set of rules of application and composition listed in section 2.2. Semantic bootstrapping models of language acquisition, such as those of Abend et al. (2017) and Mao et al.

---

[50] The present theory is therefore "pure derivational", and representationally monotonic or "nonrepresentational", in the sense of Brody (2002) and Stroik (2009:14). Since everything derivable is well-formed and there are no tranderivational constrints, it is by definition "crash-proof" in the sense of Frampton and Gutmann, 2002.)

(2021), embody those assumptions as constitutive of the Computation, using those rules directly to map (proxies for) universal and/or grounded lf to a language-specific lexicon and model of derivation. They are incapable of even representing grammars of any more powerful class, and in particular have no need to list those axioms as substantive constraints, Minimalist-style, in order for that restriction to apply.

It is often objected that this way of reducing MOVE to MERGE, unlike those involving explicit action-at-a-distance or feature-passing, requires us to forsake traditional notions of constituency, and to admit fragments like *a woman saw* and even *Alice pasta* as constituents in their own right, without the intervention of movement or deletion, as in derivations like figures 3b and 4b. Against this objection, it should be noted that the traditional constituent types are predicted by only two of the four traditional tests for constituency (lexical substitutability and ability to move), while the other two (ability to undergo coordination, and to support prosodic phrases) support exactly the superset predicted by categorial systems (Miller, 1992; Pesetsky, 1995; Steedman, 2000a; Phillips, 2003; Lechner, 2003; Jacobson, 2006).[51]

This apparent tension only arises because, faced with the phenomenon of discontinuous constituency, standard Minimalism (like most other grammar formalisms) has clung to the traditional inventory of constituent types, at the cost of introducing discontinuity into the syntactic rules themselves. By contrast, the combinatory alternative maintains strict locality in syntactic rules by giving up traditional notions of constituency and command in derivational syntax, relegating them to the level of logical form and predicate-argument structure, where they belong.[52]

References

Abbott, Barbara. 1976. Right Node Raising as a Test for Constituenthood. *Linguistic Inquiry* 7:639–642.

Abbott, Barbara. 1993. A Pragmatic Account of the Definiteness Effect in Existential Sentences. *Journal of Pragmatics* 19:39–55.

Abels, Klaus and Ad Neeleman. 2012. Linear Asymmetries and the LCA. *Syntax* 15:25–74.

Abend, Omri, Tom Kwiatkowski, Nathaniel Smith, Sharon Goldwater, and Mark Steedman. 2017. Bootstrapping Language Acquisition. *Cognition* 164:116–143.

Ades, Anthony and Mark Steedman. 1982. On the Order of Words. *Linguistics and Philosophy* 4:517–558.

Adger, David. 2003. *Core Syntax: A Minimalist Approach*. Oxford: Oxford University Press.

Adger, David. 2010. A Minimalist Theory of Feature Structure. *Features: Perspectives on a Key Notion in Linguistics*, edited by A. Kibert and Greville Corbett, 185–218. Oxford: Oxford University Press.

Adger, David. 2013. *A Syntax of Substance*. Cambridge, MA: MIT Press.

Adger, David and Gillian Ramchand. 2005. Merge and Move: *Wh*-Dependencies Revisited. *Linguistic Inquiry* 36:161–193.

Ajdukiewicz, Kazimierz. 1935. Die syntaktische Konnexität. *Polish Logic 1920–1939*, edited by Storrs McCall, 207–231. Oxford: Oxford University Press. Trans. from *Studia Philosophica*, 1, 1-27.

Asudeh, Ash, Richard Crouch, Miriam Butt, and Tracy Holloway King. 2002. Coordination and Parallelism in Glue Semantics: Integrating Discourse Cohesion and the Element Constraint. *Proceedings of the LFG02 Conference*, edited by Miriam Butt and Tracy Holloway-King, 19–39. Stanford, CA: CSLI Publications.

---

[51] The fact that *wh*-movers need to be assigned second-order categories via the language-specific lexicon explains why not everything that is a constituent can move—for example, why infinitival transitive verbs can prepose in German, but not in English—cf. note 43.

[52] The present proposal can be seen as observing a “Derivational Condition on Constituency”, as an adjunct to the standard Constituent Condition on the inputs and outputs of rules (Chomsky, 1955/1975:210-211; Chomsky, 1977:171; Steedman, 2000b:12-14).

Bach, Emmon. 1979. Control in Montague Grammar. *Linguistic Inquiry* 10:513–531.

Bach, Emmon. 1980. In Defense of Passive. *Linguistics and Philosophy* 3:297–341.

Bachrach, Asaf and Roni Katzir. 2009. Right-Node Raising and Delayed Spell-out. *InterPhases: Phase-Theoretic Investigations of Linguistic Interfaces*, edited by Kleanthes Grohmann, 283–316. Oxford: Oxford University Press.

Baldridge, Jason. 2002. *Lexically Specified Derivational Control in Combinatory Categorial Grammar*. Doctoral dissertation, University of Edinburgh.

Bar-Hillel, Yehoshua. 1953. A Quasi-Arithmetical Notation for Syntactic Description. *Language* 29:47–58.

Barker, Chris. 2012. Quantificational Binding does not Require C-Command. *Linguistic Inquiry* 43:614–633.

Barker, Chris. 2015. Scope. *The Handbook of Contemporary Semantic Theory*, edited by Lappin, Shalom and Fox, Chris, 40–76. Hoboken, NJ: John Wiley & Sons.

Barker, Chris and Chung-Chieh Shan. 2014. *Continuations and Natural Language*. Oxford University Press.

Barwise, Jon and Robin Cooper. 1981. Generalized Quantifiers and Natural Language. *Linguistics and Philosophy* 4:159–219.

Beavers, John. 2004. Type-inheritance Combinatory Categorial Grammar. *Proceedings of the International Conference on Computational Linguistics*, 57–63. Geneva: ACL.

Beavers, John and Ivan Sag. 2004a. Coordinate Ellipsis and Apparent Non-Constituent Coordination. *Proceedings of the 11th International Conference on Head-Driven Phrase Structure Grammar*, 48–69. Stanford: CSLI.

Beavers, John and Ivan Sag. 2004b. Some Arguments for Coordinate Ellipsis. *Proceedings of the 11th International Conference on Head-Driven Phrase Structure Grammar*, 48–69. Stanford,CA: CSLI Publications.

Becker, Tilman, Aravind Joshi and Owen Rambow. 1991. Long Distance Scrambling and Tree-Adjoining Grammars. *Proceedings of the Conference of the European Chapter of the Association for Computational Linguistics*, 21–26. ACL.

Bernardi, Raffaella. 2002. *Reasoning with Polarity in Categorial Type-Logic*. Doctoral dissertation, Universiteit Utrecht.

Bernardi, Raffaella. 2004. Review of Steedman 2000, *The Syntactic Process*. *Journal of Logic, Language, and Information* 13:526–530.

Berwick, Robert and Samuel Epstein. 1995a. Computational Minimalism: The Convergence of 'Minimalist' Syntax and Categorial Grammar. *Algebraic Methods in Language Processing 1995: Proceedings of the Twente Workshop on Language Technology 10, jointly held with the First Algebraic Methodology and Software Technology (AMAST) Workshop on Language Processing*, edited by Anton Nijholt, Giuseppe Scollo, and R.Steetkamp, Enschede, The Netherlands: Faculty of Computer Science, Universiteit Twente.

Berwick, Robert and Samuel Epstein. 1995b. Merge: The Categorial Imperative. *Proceedings of the First Algebraic Methodology and Software Technology (AMAST) Workshop on Language Processing*, edited by Anton Nijholt, Giuseppe Scollo, and R.Steetkamp, Enschede, The Netherlands: Faculty of Computer Science, Universiteit Twente.

Bîlbîie, Gabriela and Pegah Faghiri. 2022. An Experimental Perspective on Embedded Gapping in Persian. *The Linguistic Review* .

Boeckx, Cedric. 2008. *Bare Syntax*. Oxford: Oxford University Press.

Boeckx, Cedric. 2012. *Syntactic Islands*. Cambridge: Cambridge University Press.

Boeckx, Cedric. 2015. *Elementary Syntactic Structures: Prospects of a Feature-free Syntax*. Cambridge: Cambridge University Press.

Boeckx, Cedric, Norbert Hornstein, and Jairo Nunes. 2010. *Control as Movement*. Cambridge: Cambridge University Press.

Bolinger, Dwight. 1977. *Meaning and Form*. London: Longman.

Borer, Hagit. 1984. *Parametric Syntax: Case Studies in Semitic and Romance Languages*. Dordrecht: Foris.

Borsley, Robert, Tallerman, Maggie, and Willis, David. 2007. *The Syntax of Welsh*. Cambridge: Cambridge University Press.

Bozşahin, Cem. 2012. *Combinatory Linguistics*. Berlin: de Gruyter.

Bresnan, Joan.. 1972. *Theory of Complementation in English Syntax*. Doctoral dissertation, MIT.

Bresnan, Joan, ed.. 1982. *The Mental Representation of Grammatical Relations*. Cambridge, MA: MIT Press.

Bresnan, Joan. 2001. *Lexical-Functional Syntax*. Oxford: Blackwell.

Brody, Michael. 1995. *Lexico-Logical Form: A Radically Minimalist Theory*. Cambridge, MA: MIT Press.

Brody, Michael. 2002. On the Status of Representations and Derivations. *Derivation and Explanation in the Minimalist Program*, edited by Samuel Epstein and Daniel Seely, 19–41. Oxford: Blackwell Publishing.

Bruening, Benjamin. 2009. Selectional Asymmetries between CP and DP suggest that the DP Hypothesis is Wrong. *University of Pennsylvania Working Papers in Linguistics* 15:25–35.

Bruening, Benjamin. 2014. Precede-and-Command Revisited. *Language* 90:342–388.

Bruening, Benjamin. 2020. The Head of the Nominal is N, not D: N-to-D Movement, Hybrid Agreement, and Conventionalized Expressions. *Glossa* 5.

Bruening, Benjamin and Eman Al Khalaf. 2019. No Argument-adjunct Asymmetry in Reconstruction for Binding Condition C. *Journal of Linguistics* 55:247–276.

Büring, Daniel. 2005. *Binding Theory*. Cambridge: Cambridge University Press.

Cann, Ronnie, Ruth Kempson, and Lutz Marten. 2005. *The Dynamics of Language*. Syntax and Semantics 35. San Diego: Elsevier Academic Press.

Carlson, Greg. 1977. *Reference to Kinds in English*. Doctoral dissertation, University of Massachusetts, Amherst.

Cattell, Ray. 1976. Constraints on Movement Rules. *Language* 18–50.

Chierchia, Gennaro. 1984. *Topics in the Syntax and Semantics of Infinitives and Gerunds*. Doctoral dissertation, U. Mass. Amherst.

Chierchia, Gennaro. 1989. Anaphora and Attitudes *de se*. *Semantics and Contextual Expression*, edited by Renata Bartsch, Johan van Benthem, and Peter van Emde Boas, 1–32. Berlin: De Gruyter Mouton.

Chierchia, Gennaro. 1995. *Dynamics of Meaning*. Chicago: University of Chicago Press.

Chomsky, Noam. 1955/1975. *The Logical Structure of Linguistic Theory*. Chicago: University of Chicago Press.

Chomsky, Noam. 1977. *Essays on Form and Interpretation*. Amsterdam: North Holland.

Chomsky, Noam. 1981. *Lectures on Government and Binding*. Dordrecht: Foris.

Chomsky, Noam. 1993. A Minimalist Program for Linguistic Theory. *The View from Building 20*, edited by Kenneth Hale and Samuel Jay Keyser, 1–52. Cambridge, MA: MIT Press. Reprinted in Chomsky (1995b).

Chomsky, Noam. 1995a. Bare Phrase Structure. *Government and Binding Theory and the Minimalist Program*, edited by Gert Webelhuth, 383–439. Oxford: Blackwell.

Chomsky, Noam. 1995b. *The Minimalist Program*. Cambridge, MA: MIT Press.

Chomsky, Noam. 2000. Minimalist Inquiries: The Framework. *Step by Step: Essays in Minimalist Syntax in Honor of Howard Lasnik*, edited by Roger Martin, David Michaels, and Juan Uriagereka, 89–155. Cambridge, MA: MIT Press.

Chomsky, Noam. 2001. Derivation by Phase. *Ken Hale: a Life in Language*, edited by Michael Kenstowicz, 1–52. Cambridge, MA: MIT Press.

Chomsky, Noam. 2001/2004. Beyond Explanatory Adequacy. *Structures and Beyond: The Cartography of Syntactic Structures*, edited by Adriana Belletti, 104–131. Oxford: Oxford University Press.

Chomsky, Noam. 2007. Approaching UG from Below. *Interfaces + recursion = language? Chomsky's Minimalism and the View from Syntax-Semantics*, edited by Uli Sauerland and Hans-Martin Gärtner, 1–29. Berlin: de Gruyter.

Chomsky, Noam. 2008. On Phases. *Foundational Issues in Linguistic Theory: Essays in Honor of Jean-Roger Vergnaud*, edited by Robert Freidin, Carlos Otero, and Maria Luisa Zubizarreta, 133–166. Cambridge, MA: MIT Press.

Chomsky, Noam. 2019. The UCLA Lectures. LingBuzz.

Chomsky, Noam and Howard Lasnik. 1977. Filters and Control. *Linguistic Inquiry* 8:425–504.

Chung, Sandra. 1983. The ECP and Government in Chamorro. *Natural Language and Linguistic Theory* 1:207–244.

Cinque, Guglielmo. 1990. *Types of Ā-dependencies*. Cambridge, MA: MIT Press.

Cinque, Guglielmo. 2005. Deriving Greenberg's Universal 20 and its Exceptions. *Linguistic Inquiry* 36:315–332.

Cinque, Guglielmo. 2013. Cognition, Universal Grammar, and Typological Generalizations. *Lingua* 130:50–65.

Citko, Barbara. 2012. A Parallel Merge Solution to the Merchant/Johnson paradox. *Ways of Structure Building*, edited by Myriam Uribe-Etxebarria and Vidal Valmala, 44–65. Oxford: Oxford University Press.

Citko, Barbara. 2017. Right Node Raising. *The Wiley Blackwell Companion to Syntax*, 1–33. Hoboken, NJ: Wiley, 2 edn.

Citko, Barbara. 2018. On the Relationship between Forward and Backward Gapping. *Syntax* 21:1–36.

Citko, Barbara and Martina Gračanin-Yuksek. 2020. *Merge: Binarity in (Multidominant) Syntax*. Cambridge, MA: MIT Press.

Clark, Stephen and James Curran. 2007. Wide-Coverage Efficient Statistical Parsing with CCG and Log-Linear Models. *Computational Linguistics* 33:493–552.

Collins, Chris and Erich Groat. 2018. Distinguishing Copies and Repetitions. *Ms. Available online at* `http://ling.auf.net/lingbuzz/003809` .

Culicover, Peter W. 1993. Evidence against ECP accounts of the *that-t* effect. *Linguistic Inquiry* 557–561.

Curry, Haskell and Robert Feys. 1958. *Combinatory Logic: Vol. I*. Amsterdam: North Holland.

Davidson, Donald and Gilbert Harman, eds.. 1972. *Semantics of Natural Language*. Dordrecht: Reidel.

Davis, Anthony. 1992. Missing Heads and Empty Subjects: Underspecification in Hausa VPs. *Papers from the 28th Regional Meeting of the Chicago Linguistic Society*, 94–108. CLS.

Deal, Amy. 2009. The Origin and Content of Expletives: Evidence from "Selection". *Syntax* 12:285–323.

Den Dikken, Marcel. 1995. *Particles: On the Syntax of Verb-particle, Triadic, and Causative Constructions*. Oxford: Oxford University Press.

Diesing, Molly. 1992. *Indefinites*. Linguistic Inquiry Monograph 20. Cambridge, MA: MIT Press.

den Dikken, Marcel. 1994. Minimalist Verb (Projection) Raising. *Groninger Arbeiten zur germanistischen Linguistik* 37:71–88.

Dowty, David. 1978. Governed Transformations as Lexical Rules in a Montague Grammar. *Linguistic Inquiry* 9:393–426.

Dowty, David. 1985/1988. Type-Raising, Functional Composition, and Nonconstituent Coordination. *Categorial Grammars and Natural Language Structures*, edited by Richard Oehrle, Emmon Bach, and Deirdre Wheeler, 153–198. Dordrecht: Reidel.

Dowty, David. 2003. The Dual Analysis of Adjuncts/Complements in Categorial Grammar. *Modifying Adjuncts*, edited by Ewald Lang, Claudia Maienborn, and Cathrine Fabricius-Hansen, 33–66. Berlin: Mouton de Gruyter.

Dryer, Matthew. 1991. SVO Languages and the OV:VO Typology. *Journal of Linguistics* 27:443–482.

Eisner, Jason. 1996. Efficient Normal-Form Parsing for Combinatory Categorial Grammar. *Proceedings of the 34th Annual Meeting of the Association for Computational Linguistics, Santa Cruz, CA*, 79–86. San Francisco: Morgan Kaufmann.

Emonds, Joseph. 1979. Appositive Relatives Have No Properties. *Linguistic Inquiry* 10:211–243.

Engdahl, Elisabet. 1983. Parasitic Gaps. *Linguistics and Philosophy* 6:5–34.

Epstein, Samuel, Erich Groat, Ruriko Kawashima, and Hisatsugu Kitahara, 1998. *A Derivational Approach to Syntactic Relations*. Oxford: Oxford University Press.

Epstein, Samuel and Daniel Seely. 2006. *Derivations In Minimalism*. Cambridge: Cambridge University Press.

Farudi, Annahita. 2013. *Gapping in Farsi: A Crosslinguistic Investigation*. Doctoral dissertation, U.Mass., Amherst.

Felser, Claudia. 2004. *Wh*-copying, Phases, and Successive Cyclicity. *Lingua* 114:543–574.

Fillmore, Charles. 1965. *Indirect Object Constructions in English and the Ordering of Transformations*. The Hague: Mouton.

Fong, Sandiway. 2005. Computation with Probes and Goals. *UG and External Systems: Language, Brain and Computation*, 311–334. Amsterdam: John Benjamins.

Fox, Danny. 1999. Reconstruction, Binding Theory, and the Interpretation of Chains. *Linguistic Inquiry* 30:157–196.

Fox, Danny. 2002. Antecedent-Contained Deletion and the Copy Theory of Movement. *Linguistic Inquiry* 33:63–96.

Frampton, John and Sam Gutmann. 2002. Crash-Proof Syntax. *Derivation and Explanation in the Minimalist Program*, edited by Samuel Epstein and Daniel Seely, 90–105. Oxford: Blackwell Publishing.

Frank, Robert. 2002. *Phrase Structure Composition and Syntactic Dependencies*. Cambridge, MA: MIT Press.

Gärtner, Hans-Martin and Jens Michaelis. 2007. Some Remarks on Locality Conditions and Minimalist Grammars. *Interfaces+ Recursion= Language? Chomsky's Minimalism and the View from Syntax-Semantics*, 161–195. Berlin: Walter de Gruyter.

Gazdar, Gerald. 1981. Unbounded Dependencies and Coordinate Structure. *Linguistic Inquiry* 12:155–184.

Gazdar, Gerald, Ewan Klein, Geoffrey Pullum, and Ivan Sag. 1985. *Generalized Phrase Structure Grammar*. Oxford: Blackwell.

Geach, Peter. 1970. A Program for Syntax. *Synthèse* 22:3–17. Reprinted as Davidson and Harman 1972:483–497.

Gillies, William. 1993. Scottish Gaelic. *The Celtic Languages*, edited by Martin Ball, 145–227. London: Routledge.

Ginzburg, Jonathan and Ivan Sag. 2000. *Interrogative Investigations*. Stanford, CA: CSLI Publications.

Goldsmith, John. 1985. A Principled Exception to the Coordinate Structure Constraint. *Proceedings of the 21st Regional Meeting of the Chicago Linguistic Society*. Chicago: CLS.

Goodall, Grant. 1987. *Parallel Structures in Syntax*. Cambridge: Cambridge University Press.

Haegeman, Liliane and Henk van Riemsdijk. 1986. Verb Projection Raising, Scope, and the Typology of Rules Affecting Verbs. *Linguistic Inquiry* 17:417–466.

Haider, Hubert. 2003. V-clustering and Clause Union: Causes and Effects. *Verb Constructions in German and Dutch*, edited by Pieter Seuren and Gerhard Kempen, 91–126. Amsterdam: John Benjamins.

Hale, Kenneth and Samuel Jay Keyser. 2002. *Prolegomenon to a Theory of Argument Structure*. Linguistic Inquiry Monograph 39. Cambridge, MA: MIT Press.

Hankamer, Jorge. 1971. *Constraints on Deletion in Syntax*. Doctoral dissertation, Yale University. Published as Hankamer (1979).

Hankamer, Jorge. 1979. *Deletion in Coordinate Structures*. New York: Garland. Hankamer (1971), revised.

Hankamer, Jorge and Ivan Sag. 1976. Deep and Surface Anaphora. *Linguistic Inquiry* 7:391–428.

Hardt, Daniel and Maribel Romero. 2004. Ellipsis and the Structure of Discourse. *Journal of Semantics* 21:375–414.

Harkema, Hendrik. 2001. *Parsing Minimalist Languages*. Doctoral dissertation, University of California Los Angeles.

Hawkins, John. 1982. Cross Category Harmony, $\bar{X}$ and the Predictions of Markedness. *Journal of Linguistics* 18:1–35.

Heim, Irene and Angelika Kratzer. 1998. *Semantics in Generative Grammar*. Oxford: Blackwell.

Hepple, Mark. 1990. *The Grammar and Processing of Order and Dependency: A Categorial Approach*. Doctoral dissertation, University of Edinburgh.

Heycock, Caroline. 1995. Asymmetries in Reconstruction. *Linguistic Inquiry* 547–570.

Higginbotham, James. 1983. Logical Form, Binding, and Nominals. *Linguistic inquiry* 14:395–420.

Hirsch, Aron. 2017. A Case for Conjunction Reduction. *submitted* .

Hirsch, Aron and Michael Wagner. 2015. Right Node Raising, Scope, and Plurality. *Proceedings of the 20th Amsterdam Colloquium*, 187–196.

Hockenmaier, Julia. 2003. *Data and Models for Statistical Parsing with CCG*. Doctoral dissertation, School of Informatics, University of Edinburgh.

Hockenmaier, Julia and Peter Young. 2008. Non-local Scrambling: The Equivalence of TAG and CCG Revisited. *Proceedings of the 9th International Workshop on Tree-Adjoining Grammars and Related Formalisms (TAG+9)*. Tübingen.

Hornstein, Norbert. 1995. *Logical Form: From GB to Minimalism*. Oxford: Blackwell.

Hornstein, Norbert. 1999. Movement and Control. *Linguistic Inquiry* 30:69–96.

Hornstein, Norbert. 2001. *Move! A Minimalist Theory of Construal*. Oxford: Blackwell Publishing.

Hornstein, Norbert and Juan Uriagereka. 2002. Reprojections. *Derivation and Explanation in the Minimalist Program* edited by Samuel Epstein and Daniel Seely, 106–132. Blackwell Publishing.

Huang, C.-T. James. 1982. *Logical Relations in Chinese and the Theory of Grammar*. Doctoral dissertation, MIT.

Huddleston, Rodney and Geoffrey Pullum. 2002. *Cambridge Grammar of the English Language*. Cambridge: Cambridge University Press.

Hudson, Richard. 1976. Conjunction Reduction, Gapping, and Right-Node Raising. *Language* 535–562.

Jackendoff, Ray. 1972. *Semantic Interpretation in Generative Grammar*. Cambridge, MA: MIT Press.

Jackendoff, Ray. 1997. *The Architecture of the Language Faculty*. Linguistic Inquiry Monograph 28. Cambridge, MA: MIT Press.

Jacobson, Pauline. 1990. Raising as Function Composition. *Linguistics and Philosophy* 13:423–476.

Jacobson, Pauline. 1994. Binding Connectivity in Copular Sentences. *Semantics and Linguistic Theory*, vol. 4, 161–178. Cornell University.

Jacobson, Pauline. 1999. Towards a Variable-Free Semantics. *Linguistics and Philosophy* 22:117–184.

Jacobson, Pauline. 2006. Constituent Structure. *The Encyclopedia of Language and Linguistics*, vol. 3, edited by Keith Brown, 58–71. Amsterdam: Elsevier, 2 edn.

Jacobson, Pauline. 2014. *Compositional Semantics*. Oxford: Oxford University Press.

Johnson, Kyle. 2017. Gapping. *The Wiley Blackwell Companion to Syntax*, edited by Martin Everaert and Henk van Riemsdijk, 1–40. Hoboken, NJ: Wiley Online Library, 2nd edn.

Johnson, Kyle. 2020. Principle C. *submitted* .

Joshi, Aravind. 1985. Tree-Adjoining Grammars: How much Context-Sensitivity id Required to Provide Sensible Structural Descriptions? *Natural Language Parsing*, edited by David Dowty, Lauri Karttunen, and Arnold Zwicky, 206–250. Cambridge: Cambridge University Press.

Joshi, Aravind, K.Vijay-Shanker, and David Weir. 1991. The Convergence of Mildly Context-Sensitive Formalisms. *Processing of Linguistic Structure*, edited by Peter Sells, Stuart Shieber, and Tom Wasow, 31–81. Cambridge, MA: MIT Press.

Jurka, Johannes. 2013. Subject Islands in German Revisited. *Experimental Syntax and Island Effects*, edited by Jon Sprouse and Norbert Hornstein, 265–285. Cambridge: Cambridge University Press.

Kamp, Hans. 1981/1984. A Theory of Truth and Semantic Representation. *Truth, Interpretation, and Information*, edited by Jeroen Groenendijk, Theo Janssen, and Martin Stokhof, 1–41. Dordrecht: Foris.

Kay, Martin. 1984. Functional Unification Grammar: A Framework for Machine Translation. *Proceedings of the 22nd Annual Conference of the Association for Computational Linguistics, Stanford, CA*, 75–78. ACL Anthology.

Kayne, Richard. 1983. *Connectedness and Binary Branching*. Dordrecht: Foris.

Kayne, Richard. 1994. *The Antisymmetry of Syntax*. Cambridge, MA: MIT Press.

Kehler, Andrew. 2002. *Coherence, Reference, and the Theory of Grammar*. Stanford, CA: CSLI Publications.

Keine, Stefan and Rajesh Bhatt. 2016. Interpreting Verb Clusters. *Natural Language & Linguistic Theory* 34:1445–1492.

Keine, Stefan and Ethan Poole. 2017. Intervention in *tough*-Constructions Revisited. *The Linguistic Review* 34:295–329.

Keine, Stefan and Ethan Poole. 2018. Not all Reconstruction Effects are Syntactic. *Ms., University of Southern California and University of California, Los Angeles* .

Kobele, Gregory. 2010. Without Remnant Movement, MGs are Context-Free. *Proceedings of 11th Conference on the Mathematics of Language (MoL)*, 160–173. Berlin: Springer.

Koller, Alexander and Marco Kuhlmann. 2009. Dependency Trees and the Strong Generative Capacity of CCG. *Proceedings of the 12th Conference of the European Chapter of the ACL*, 460–468. Athens: ACL.

Koopman, Hilda. 1996. The Spec Head Configuration. *Syntax at sunset: UCLA working papers in syntax and semantics* 1.

Koopman, Hilda and Anna Szabolcsi. 2000. *Verbal Complexes*. Cambridge, MA: MIT Press.

Koster, Jan. 1978. *Locality Principles in Syntax*. Dordrecht: Foris.

Kratzer, Angelika. 1988/1995. Stage-Level and Individual-Level Predicates. *The Generic Book*, edited by Gregory Carlson and Geoffrey Pelletier, 125–175. Chicago: Chicago University Press.

Kubota, Yusuke and Robert Levine. 2020. *Type-Logical Syntax*. Cambridge, MA:MIT Press.

Kuhlmann, Marco, Alexander Koller, and Giorgio Satta. 2010. The Importance of Rule Restrictions in CCG. *Proceedings of the 48th Annual Meeting of the Association for Computational Linguistics*, 534–543. Uppsala: ACL.

Kuhlmann, Marco, Alexander Koller, and Giorgio Satta. 2015. Lexicalization and Generative Power in CCG. *Computational Linguistics* 41:187–219.

Kuno, Susumu. 1987. *Functional Syntax: Anaphora, Discourse, and Empathy*. Chicago: University of Chicago Press.

Lakoff, George. 1986. Frame Semantic Control of the Coordinate Structure Constraint. *Papers from the Parasession on Pragmatics and Grammatical Theory at the 22nd Regional Meeting of the Chicago Linguistics Society*, edited by Anne Farley, Peter Farley, and Karl-Eric McCullough, 152–167. Chicago: CLS.

Lambek, Joachim. 1958. The Mathematics of Sentence Structure. *American Mathematical Monthly* 65:154–170.

Landau, Idan. 2001. *Elements of Control: Structure and Meaning in Infinitival Constructions*. Berlin: Springer.

Landau, Idan. 2015. *A Two-Tiered Theory of Control*. Cambridge, MA: MIT Press.

Landau, Idan. 2021. *A Selectional Theory of Adjunct Control*. Cambridge, MA: MIT Press.

Lasnik, Howard. 2001. Subjects, Objects, and the EPP. *Objects and other Subjects*, edited by William Davies and Stanley Dubinsky, 103–121. Berlin: Springer.

Lebeaux, David. 1991. Relative Clauses, Licensing, and the Nature of the Derivation. *Syntax and Semantics 25: Perspectives on Phrase Structure: Heads and Licensing*, 209–239. San Diego: Academic Press.

Lebeaux, David. 2009. *Where does Binding Theory Apply?* Cambridge, MA: Mit Press.

Lechner, Winfried. 1998. Two Kinds of Reconstruction. *Studia Linguistica* 52:276–310.

Lechner, Winfried. 2003. Phrase Structure Paradoxes, Movement and Ellipsis. *The Interfaces: Deriving and Interpreting Omitted Structures*, edited by Kerstin Schwabe and Susanne Winkler, 177–203. Amsterdam: John Benjamins.

Lechner, Winfried. 2013. Diagnosing Covert Movement: The Duke of York and Reconstruction. *Diagnosing Syntax*, edited by Lisa Cheng and Norbert Corver, 158–189. Oxford: Oxford University Press.

Lechner, Winfried. 2018. A Calculus for Reconstruction and Anti-Reconstruction. *Reconstruction Effects in Relative Clauses*, edited by Manfred Krifka and Mathias Schenner, 113–144. Berlin: De Gruyter.

Legate, Julie Anne. 1999. The Morphosyntax of Irish Agreement. *MIT Working Papers in Linguistics* 33:219–240.

Levin, Beth. 1993. *English Verb-Classes and Alternations: A Preliminary Study*. Chicago: University of Chicago Press.

Levine, Robert. 2017. *Syntactic Analysis: An HPSG-based Approach*. Cambridge: Cambridge University Press.

Lewis, David. 1970. General Semantics. *Synthèse* 22:18–67. Reprinted as Ch.1 of Partee (1976).

Lewis, David. 1979a. Attitudes de dicto and de se. *The philosophical review* 88:513–543.

Lewis, David. 1979b. Scorekeeping in a Language Game. *Journal of Philosophical Logic* 8:339–359.

Lewis, Mike and Steedman, Mark. 2014. A^* CCG Parsing with a Supertag-factored Model. *Proceedings of the Conference on Empirical Methods in Natural Language Processing*, 990–1000. Doha, Qatar: ACL.

Li, Charles and Sandra Thompson. 1975. The Semantic Function of Word Order: A Case Study in Mandarin. *Word Order and Word Order Change*, edited by Charles Li, 163–195. Austin, TX: University of Texas Press.

Lidz, Jeffrey. 2018. The Scope of Children's Scope: Representation, Parsing and Learning. *Glossa* 3.

Lipták, Anikó. 2019. Hungarian. *The Oxford Handbook of Ellipsis*, edited by Jeroen Van Craenenbroeck and Tanja Temmerman, 815–840. Oxford: Oxford University Press.

Maling, Joan. and Annie Zaenen. 1978. Nonuniversality of a Surface Filter. *Linguistic Inquiry* 9:475–497.

Man, Haixia and Mark Steedman. 2023. Word-order and Eventuality in Chinese. *in preparation* .

Manzini, Rita. 1992. *Locality: A Theory and some of its Empirical Consequences*. Cambridge, MA: MIT Press.

Mao, Jiayuan, Freda Shi, Jiajun Wu, Roger Levy, and Joshua Tenenbaum. 2021. Grammar-Based Grounded Lexicon Learning. *Proceedings of the Thirty-Fifth Conference on Neural Information Processing Systems*, 7865–7878.

Marantz, Alec. 1997. No Escape from Syntax: Don't Try Morphological Analysis in the Privacy of your own Lexicon. *Proceedings of the 21st Penn Linguistics Colloquium*, *University of Pennsylvania Working Papers in Linguistics*, vol. 4, 201–225.

McCawley, James. 1982. Parentheticals and Discontinuous Constituent Structure. *Linguistic Inquiry* 13:91–106.

McCawley, James. 1987. Some Additional Evidence for Discontinuity. *Discontinuous Constituency*, edited by Almerindo Ojeda and Geoffrey Huck, Syntax and Semantics 20, 185–200. New York: Academic Press.

McCloskey, James. 1979. *Transformational Syntax and Model Theoretic Semantics: A Case Study in Modern Irish*. Dordrecht: Reidel.

McCloskey, James. 1986. Right Node Raising and Preposition Stranding. *Linguistic Inquiry* 17:183–186.

McCloskey, James. 1990. Resumptive Pronouns, Ā-Binding, and Levels of Representations in Irish. *The Syntax of the Modern Celtic Languages*, *Syntax and Semantics*, vol. 23, edited by Randall Hendrick, 199–248. New York: Academic Press.

McCloskey, James. 2017. New Thoughts on Old Questions–Resumption in Irish. *Asking the Right Question: Essays in Honor of Sandra Chung*, edited by James Ostrove, Ruth Kramer, and Joseph Sabbagh, 81–102. Santa Cruz, CA: University of California.

Miller, Philip. 1992. *Clitics and Constituents in Phrase Structure Grammar*. New York: Garland.

Miller, Philip. 1999. *Strong Generative Capacity: The Semantics of Linguistic Formalism*. Stanford, CA: CSLI publications.

Montague, Richard. 1970. English as a Formal Language. *Linguaggi nella Società e nella Technica*, edited by Bruno Visentini, 189–224. Milan: Edizioni di Communità. Reprinted as Thomason 1974:188-221.

Montague, Richard. 1973. The Proper Treatment of Quantification in Ordinary English. *Approaches to Natural Language: Proceedings of the 1970 Stanford Workshop on Grammar and Semantics*, edited by Jaakko Hintikka, Julius Moravcsik, and Patrick Suppes, 221–242. Dordrecht: Reidel. Reprinted in Thomason 1974:247-279.

Moortgat, Michael. 1988. *Categorial Investigations*. Doctoral dissertation, Universiteit van Amsterdam. Published by Foris, Dordrecht, 1989.

Morrill, Glynn. 2011. *Categorial Grammar: Logical Syntax, Semantics, and Processing*. Oxford: Oxford University Press.

Moulton, Keir and Chung-Hye Han. 2018. C-command vs. Scope: An Experimental Assessment of Bound-Variable Pronouns. *Language* 94:191–219.

Munn, Alan. 1999. On the Identity Requirement of ATB Extraction. *Natural Language Semantics* 7:421–425.

Neeleman, Ad and Hans van de Koot. 2010. A Local Encoding of Syntactic Dependencies and Its Consequences for the Theory of Movement. *Syntax* 13:331–372.

Neijt, Anneke. 1979. *Gapping*. Dordrecht: Foris.

Nunes, Jairo. 2004. *Linearization of Chains and Sideward Movement*. Linguistic Inquiry Monograph 43. Cambridge, MA: MIT Press.

Oehrle, Richard. 1976. *The Grammatical Status of the English Dative Alternation*. Doctoral dissertation, MIT.

Ojeda, Almerindo and Geoffrey Huck, eds.. 1987. *Discontinuous Constituency*. Syntax and Semantics 20. New York: Academic Press.

Otsuka, Yuko. 2005. Scrambling and Information Focus: VSO-VOS Alternation in Tongan. *The Free Word Order Phenomenon: Its Syntactic Sources and Diversity*, edited by Joachim Sabel and Mamoru Saito, 243–280. Berlin: Mouton de Gruyter.

Partee, Barbara, ed.. 1976. *Montague Grammar*. New York: Academic Press.

Partee, Barbara and Mats Rooth. 1983. Generalised Conjunction and Type Ambiguity. *Meaning, Use, and Interpretation of Language*, edited by Rainer Bäuerle, Christoph Schwarze, and Arnim von Stechow, 361–383. Berlin: de Gruyter.

Pesetsky, David. 1982. *Paths and Categories*. Doctoral dissertation, MIT.

Pesetsky, David. 1995. *Zero Syntax*. Cambridge, MA: MIT Press.

Peters, Stanley and Robert Ritchie. 1973. On the Generative Power of Transformational Grammars. *Information Science* 6:49–83.

Phillips, Colin. 2003. Linear Order and Constituency. *Linguistic Inquiry* 34:37–90.

Pierrehumbert, Janet and Julia Hirschberg. 1990. The Meaning of Intonational Contours in the Interpretation of Discourse. *Intentions in Communication*, edited by Philip Cohen, Jerry Morgan, and Martha Pollack, 271–312. Cambridge, MA: MIT Press.

Polinsky, Maria, Carlos Gallo, Peter Graff, Ekaterina Kravtchenko, Adam Milton Morgan, and Anne Sturgeon. 2013. Subject Islands are Different. *Experimental Syntax and Island Effects*, edited by Jon Sprouse and Norbert Hornstein, 286–309. Cambridge: Cambridge University Press.

Pollard, Carl and Ivan Sag. 1994. *Head-Driven Phrase Structure Grammar*. Stanford, CA: CSLI Publications.

Postal, Paul. 1966. On So-called Pronouns in English. *Report of the 17th Annual Round Table on Linguistics and Language Studies*, edited by Francis Dinneen, 177–206. Wahington, DC: Georgetown University Press.

Postal, Paul. 1974. *On Raising: One Rule of English Grammar and its Theoretical Implications*. Cambridge, MA: MIT Press.

Postal, Paul. 1994. Parasitic and Pseudoparasitic Gaps. *Linguistic inquiry* 25:63–117.

Postal, Paul. 1998. *Three Investigations of Extraction*. Cambridge, MA: MIT Press.

Prince, Ellen. 1981. Towards a Taxonomy of Given-New Information. *Radical Pragmatics*, edited by Peter Cole, 223–256. New York: Academic Press.

Pullum, Geoffrey and Philip Miller. 2022. NPs vs DPs: Why Chomsky was Right. LingBuzz/006845.

Radford, Andrew. 2004. *Minimalist Syntax: Exploring the Structure of English*. Cambridge: Cambridge University Press.

Rando, Emily and Donna Jo Napoli. 1978. Definites in *There*-Sentences. *Language* 300–313.

Reinhart, Tanya. 1983. *Anaphora and Semantic Interpretation*. London: Croom Helm.

Reinhart, Tanya. 1997. Quantifier Scope: How Labor is Divided between QR and Choice Functions. *Linguistics and philosophy* 335–397.

Reinhart, Tanya. 2006. *Interface Strategies: Optimal and Costly Computations*. Linguistic Inquiry Monograph 45. Cambridge, MA: MIT Press.

Reinhart, Tanya and Eric Reuland. 1991. Anaphors and Logophors: An Argument Structure Perspective. *Long Distance Anaphora*, edited by Jan Koster and Eric Reuland, 283–322. Cambridge: Cambridge University Press.

Rizzi, Luigi. 1990. *Relativized Minimality*. Cambridge, MA: MIT Press.

Rogers, James. 1998. *A Descriptive Approach to Language-Theoretic Complexity*. Stanford, CA: CSLI Publications.

Rogers, James. 2003. wMSO Theories as Grammar Formalisms. *Theoretical Computer Science* 293:291–320.

Romero, Maribel. 1998. *Focus and Reconstruction Effects in* wh-*phrases*. Doctoral dissertation, University of Massachusetts Amherst.

Rosenbaum, Peter. 1967. *The Grammar of English Predicate Complement Constructions*. Cambridge, MA: MIT Press.

Ross, John Robert. 1967. *Constraints on Variables in Syntax*. Doctoral dissertation, MIT. Published as Ross 1986.

Ross, John Robert. 1970. Gapping and the Order of Constituents. *Progress in Linguistics*, edited by Manfred Bierwisch and Karl Heidolph, 249–259. The Hague: Mouton.

Ross, John Robert. 1986. *Infinite Syntax!* Norton, NJ: Ablex.

Rothstein, Susan. 2001. *Predicates and their Subjects*. Dordrecht: Kluwer.

Sabbagh, Joseph. 2007. Ordering and Linearizing Rightward Movement. *Natural Language & Linguistic Theory* 25:349–401.

Sag, Ivan. 1982. A Semantic Theory of "NP-Movement" Dependencies. *The Nature of Syntactic Representation*, edited by Pauline Jacobson and Geoffrey Pullum, 427–466. Dordrecht: Reidel.

Sag, Ivan. 1997. English Relative Clause Constructions. *Journal of linguistics* 33:431–483.

Sag, Ivan, Gerald, Thomas Wasow, and Steven Weisler. 1985. Coordination and How to Distinguish Categories. *Natural Language and Linguistic Theory* 3:117–172.

Sauerland, Uli and Fabian Heck. 2003. LF-Intervention Effects in Pied-piping. *Proceedings of the 33rd Conference of the North Eastern Linguistics Society (NELS)*, 347–366.

Schiffer, Lena Katharina and Andreas Maletti. 2021. Strong Equivalence of TAG and CCG. *Transactions of the Association for Computational Linguistics* 9:707–720.

Selkirk, Elisabeth. 2011. The Syntax-Phonology Interface. *The Handbook of Phonological Theory*, edited by John Goldsmith, Jason Riggle, and Alan Yu, 435–483. Oxford: Blackwell, 2nd edn.

Shieber, Stuart. 1985. Evidence against the Context-Freeness of Natural Language. *Linguistics and Philosophy* 8:333–343.

Shieber, Stuart. 1986. *An Introduction to Unification-Based Approaches to Grammar*. Stanford, CA: CSLI Publications.

Smith, Neil and Annabel Cormack. 2015. Features from *Aspects* via the Minimalist Program to Combinatory Categorial Grammar. *50 Years Later: Reflections on Chomsky's* Aspects, edited by Ángel Gallego and Dennis Ott, 233–247. MIT Working Papers in Linguistics.

Sportiche, Dominique. 2019. Somber Prospects for Late Merger. *Linguistic Inquiry* 50:416–424.

Stalnaker, Robert. 1974. Pragmatic Presuppositions. *Semantics and Philosophy*, edited by Milton Munitz and Peter Unger, 197–214. New York: New York University Press.

Stanojević, Miloš and Mark Steedman. 2021. Formal Basis of a Language Universal. *Computational Linguistics* 47:9–42.

Starke, Michal. 2001. *Move Dissolves into Merge: a Theory of Locality*. Doctoral dissertation, Université de Genève.

Starke, Michal. 2004. On the Inexistence of Specifiers and the Nature of Heads. LingBuzz

von Stechow, Arnim. 1991. Syntax und Semantik. *Semantik: Ein Handbuch der zeitgenössischen Forschung/Semantics: An International Handbook of Contemprorary Research*, edited by Arnim von Stechow and Dieter Wunderlich, 90–148. Berlin: de Gruyter.

von Stechow, Arnim. 1996. Against LF Pied-Piping. *Natural Language Semantics* 4:57–110.

Steedman, Mark. 1985. Dependency and Coordination in the Grammar of Dutch and English. *Language* 61:523–568.

Steedman, Mark. 1985/1988. Combinators and Grammars. *Categorial Grammars and Natural Language Structures*, edited by Richard Oehrle, Emmon Bach, and Deirdre Wheeler, 417–442. Dordrecht: Reidel.

Steedman, Mark. 1987. Combinatory Grammars and Parasitic Gaps. *Natural Language and Linguistic Theory* 5:403–439.

Steedman, Mark. 1990. Gapping as Constituent Coordination. *Linguistics and Philosophy* 13:207–263.

Steedman, Mark. 1991a. Structure and Intonation. *Language* 67:262–296.

Steedman, Mark. 1991b. Type-Raising and Directionality in Combinatory Grammar. *Proceedings of the 29th Annual Meeting of the Association for Computational Linguistics*, 71–78. Berkeley, CA: ACL.

Steedman, Mark. 1996. *Surface Structure and Interpretation*. Linguistic Inquiry Monograph 30. Cambridge, MA: MIT Press.

Steedman, Mark. 1999. Quantifier Scope Alternation in CCG. *Proceedings of the 37th Annual Meeting of the Association for Computational Linguistics*, 301–308. College Park, MD: ACL.

Steedman, Mark. 2000a. Information Structure and the Syntax-Phonology Interface. *Linguistic Inquiry* 34:649–689.

Steedman, Mark. 2000b. *The Syntactic Process*. Cambridge, MA: MIT Press.

Steedman, Mark. 2002. Plans, Affordances, and Combinatory Grammar. *Linguistics and Philosophy* 25:723–753.

Steedman, Mark. 2012. *Taking Scope: The Natural Semantics of Quantifiers*. Cambridge, MA: MIT Press.

Steedman, Mark. 2014a. Evolutionary Basis for Human Language: Comment on "Toward a Computational Framework for Cognitive Biology: Unifying Approaches from Cognitive Neuroscience and Comparative Cognition" by Tecumseh Fitch. *Physics of Life Reviews* 11:382–388.

Steedman, Mark. 2014b. The Surface-Compositional Semantics of English Intonation. *Language* 90:2–57.

Steedman, Mark. 2018. The Lost Combinator. *Computational Linguistics* 44:613–629.

Steedman, Mark. 2020. A Formal Universal of Natural Language Grammar. *Language* 96:618–660.

Stroik, Thomas. 2009. *Locality in Minimalist Syntax*. Cambridge, MA: MIT Press.

Szabolcsi, Anna. 1983. ECP in Categorial Grammar. Ms., Max Planck Institute, Nijmegen.

Szabolcsi, Anna. 1983/1989. Bound Variables in Syntax: Are There Any? *Semantics and Contextual Expression*, edited by Renate Bartsch, Johan van Benthem, and Peter van Emde Boas, 295–318. Dordrecht: Foris.

Szabolcsi, Anna. 1983/1992. Combinatory Grammar and Projection from the Lexicon. *Lexical Matters*, edited by Ivan Sag and Anna Szabolcsi, 241–268. Stanford, CA: CSLI Publications.

Szabolcsi, Anna. 2007. Strong vs. Weak Islands. *The Blackwell Companion to Syntax*, edited by Martin Everaert and Henk van Riemsdijk, 479–531. Oxford: Wiley-Blackwell.

Szabolcsi, Anna and Frans Zwarts. 1993. Weak Islands and an Algebraic Semantics for Scope Taking. *Natural Language Semantics* 1:235–284.

Takahashi, Shoichi and Sarah Hulsey. 2009. Wholesale Late Merger: Beyond The A/Ā Distinction. *Linguistic Inquiry* 40:387–426.

Thomason, Richmond, ed.. 1974. *Formal Philosophy: Papers of Richard Montague*. New Haven, CT: Yale University Press.

Torr, John. 2019. *Broad-Coverage Parsing with Minimalist Grammars*. Doctoral dissertation, University of Edinburgh.

Torr, John, Miloč Stanojević, Mark Steedman, and Shay Cohen. 2019. Wide-Coverage Neural A* Parsing for Minimalist Grammars. *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, 2486–2505. ACL.

Truswell, Robert. 2007a. Extraction from Adjuncts and the Structure of Events. *Lingua* 117:1355–1377.

Truswell, Robert. 2007b. *Locality of Wh-Movement and the Individuation of Events*. Doctoral dissertation, University College London.

Uchida, Hiroyuki. 2008. *Scope alternation and type logical grammar*. Doctoral dissertation, University of London, University College London (United Kingdom).

VanLehn, Kurt. 1978. *Determining the Scope of English Quantifiers*. Master's thesis, MIT, Cambridge, MA. AI-TR-483, Artificial Intelligence Laboratory.

Vennemann, Theo and Ray Harlow. 1977. Categorial Grammar and Consistent Basic VX Serialization. *Theoretical Linguistics* 4:227–254.

Vergnaud, Jean Roger. 1977/2006. Letter to Noam Chomsky and Howard Lasnik on "Filters and Control," April 17, 1977. *Syntax: Critical Concepts in Linguistics*, edited by Robert Freidin and Howard Lasnik, 21–34. Oxford: Blackwell.

Wallenberg, Joel. 2009. *Antisymmetry and the Conservation of c-Command: Scrambling and Phrase Structure in Synchronic and Diachronic Perspective*. Doctoral dissertation, University of Pennsylvania.

Wallenberg, Joel. 2013. Scrambling, LF, and Phrase Structure Change in Yiddish. *Lingua* 133:289–318.

Wasow, Thomas. 1972. *Anaphoric Relations in English*. Doctoral dissertation, Massachusetts Institute of Technology, Cambridge, MA. Published with extensive revisions as Wasow (1979).

Wasow, Thomas. 1979. *Anaphora in Generative Grammar*. Ghent: E.Story-Scientia.

Weir, David. 1988. *Characterizing Mildly Context-Sensitive Grammar Formalisms*. Doctoral dissertation, University of Pennsylvania, Philadelphia. Published as Technical Report CIS-88-74.

Wells, Rulon. 1947. Immediate Constituents. *Language* 23:81–117.

Wexler, Kenneth and Peter Culicover. 1980. *Formal Principles of Language Acquisition*. Cambridge, MA: MIT Press.

Wilder, Chris. 1999. Right Node Raising and the LCA. *Proceedings of 18th WCCFL*, 586–598. Somerville MA: Cascadilla Press.

Williams, Edwin. 1978. Across-the-Board Rule Application. *Linguistic Inquiry* 9:31–43.

Williams, Edwin. 1984. *There*-insertion. *Linguistic inquiry* 131–153.

Williams, Edwin. 1994. *Thematic Structure in Syntax*. Cambridge, MA: MIT Press.

Williams, Edwin. 2003. *Representation Theory*. Cambridge, MA: MIT Press.

Woods, William. 1968a. Procedural Semantics for a Question-Answering Machine. *Proceedings of the Fall Joint Computer Conference, part I*, 457–471. ACM.

Woods, William. 1968b. *Semantics for a Question-Answering System*. Doctoral dissertation, Harvard. Tech Report NSF-19, Harvard Computing Laboratory.

Woods, William. 1970. Transition Network Grammars for Natural Language Analysis. *Communications of the Association for Computing Machinery* 18:264–274.

Wurmbrand, Susi. 2004. West Germanic Verb Clusters: The Empirical Domain. *Verb clusters: A study of Hungarian, German, and Dutch*, edited by Katalin É. Kiss and Henk van Riemsdijk, 43–85. Amsterdam: John Benjamins.

Zhang, Niina. 2010. *Coordination in Syntax*. Cambridge: Cambridge University Press.

Mark Steedman
School of Informatics
University of Edinburgh
10 Crichton Street
Edinburgh EH8 9AB
United Kingdom
`steedman@inf.ed.ac.uk`