## Chapter 9

## Wh-Constructions

Le silence éternel de ces espaces infinis m'effraie. ${ }^{1}$
-Pensées, I[ 206 Blaise Pascal, 1671

The preceding chapters have argued that all bounded constructions, including those like raising and there-insertion that allow dependencies to grow across unboundedly many non-finite clauyse boundaries, are lexically headed.
The unbounded constructions such as relativization are those in which the elements of a semantic dependency are separated by unboundedly many finiteclause boundaries without being semantically arguments or dependents of the intervening verbs, as the long-distance raised/controlled arguments are in (??) and (??) of chapter 9 . They have widely been held to require rules of discontinuity or "action at a distance, such as $w h$-movement. The present chapter argues that they too are lexically headed, this time to a first approximation by the $w h$-item.
The account proposed here adheres to the following principle, which further constrains the lexicon to which combinatory projection applies: ${ }^{2}$

## (1) The Projective Dependency Principle (PDP)

A single non-disjunctive lexical category for the governor of a given construction specifies both the bounded dependencies that arise when its complements are in canonical position and the unbounded dependencies that arise when those complements are displaced under relativization, coordination, and the like.

The latter principle does not prevent a given word from heading more than one construction, and hence being associated with more than one category. Nor does it exclude the possibility that a given word-sense pair may permit more than one canonical order, and hence have more than one category per sense. The claim is simply that, normally, a single category specifies both local and long-range syntactic and semantic dependencies for the construction

[^0]that it governs. It follows that the dependencies that in the terms of Dependency Grammar (Hays, 1964, 1967) are called "non-projective" cannot exist: all dependency is projective.

For example, we shall see in later sections that the single lexical syntactic category $V P / N P: \lambda x \lambda y$.see $x y$ that we have assumed for the transitive verb "see", which simply specifies an object to its right and does not distinguish between "antecedent," " $\theta$," or any other variety of government, is involved in all of the dependencies underlined in (2):
(2) a. Gabbitas sees a bird.
b. Gabbitas sees every bird.
c. the bird that I believe that Gabbitas sees
d. I believe that Gabbitas sees and you believe that Thring hears, a bird.

The principle (1) has the effect of keeping the grammar relatively small. However, just as the definition above allows the exceptional possibility that an argument only appears in situ, it also allows categories to specify arguments as exceptionally only supporting extraction. The impossibility in general of subordinate subject extraction in rigid SVO languages like English, and the existence of a small class of English bare-complement verbs like believe that do allow extraction of embedded subjects means that those verbs require an additional special case lexical category to license the extraction. These and other cases are discussed in Steedman 1991, and more briefly in section 9.6 of chapter 9.

Like the Minimalist assumption of the Copy Theory of Movement, the present Projective Dependency Principle (1) makes redundant Chomsky's 1981 Empty Category Principle (ECP), which said that the trace residues of movement had to be in positions where the verb involved would normally govern its arguments.

The Projective Dependency Principle (1) distinguishes CCG and the Minimalist Program from both TAG and GPSG, in which local and long-distance dependencies are mediated by different initial trees, rules, and/or categories, and from HPSG, in which they are mediated by a disjunctive category using SLASH features. ${ }^{3}$
3. HPSG SLASH features are a descendant of GPSG's slash notation for passing long-range dependency information through the CF derivation. While there is a historical reason for the GPSG/HPSG slash features being so named (Gazdar 1981:159), they are, as noted earlier, interpreted quite differently from categorial slashes, as specifically denoting extraction, rather than general-purpose specification. LFG represents extraction by "functional uncertainty", defining an f -structural path between source and target of what is in other respects a movement analysis.

It is the inclusion of the rules of composition introduced as (6) in chapter 4.1 that will allow capture of relativization and other unbounded phenomena in CCG without resorting to movement or any related notion of "action at a distance". The first such construction we will consider is the Wh-question.

### 9.1 Wh-questions

The following observations concerning the formation of Wh-questions by Koopman (1983) can be explained in terms of the earlier discussion of SubjectAUX inversion in section 7.1 of the last chapter and the categories shown for the $W h$-element. EnglishObject wh-questions require subject-aux-inversion, as in (3), and are not equivalent to noninverting object indirect questions like (15): ${ }^{4}$
(3)
(4)

| What | Mary bought |
| :---: | :---: |
| $S_{w h / o t}(S / N P)$ | $N P^{\uparrow} \quad V P / N P$ |
|  | $S / N P$ |
| $S_{w h}($ | $\pm S_{w h q}$ ) |

Subject wh-questions require the non-inverting verb category, and exclude dosupport:
(5) $\frac{\text { Who }}{\frac{S_{w h q / \diamond \star}(S \backslash N P)}{\text { bought Ulysses? }}} \underset{S_{w h q}}{S \backslash N P}$
(6) $\frac{* \text { Who }}{S_{\text {whq } / \text { * }}(S \backslash N P)} \frac{\text { did }}{\left(\overline{\left.S_{i n v} / V P\right) / N P}\right.} \frac{\text { buy Ulysses? }}{V P}$
4. We suppress details of the semantic side of the derivation until the discussion of the related relative clause construction in section 9.4.
(The latter is ungrammatical in the absence of contrastive accent on did.) ${ }^{5}$
Wh-determiners project these categories onto more complex wh-items:
(7)

| What | book | did Mary | buy? |
| :---: | :---: | :---: | :---: |
| $\overline{\left(S_{\text {whq/ }}^{\text {ct }}\right.}$ ( $\left.\left.S_{\text {inv }} / N P\right)\right) / N$ | $N$ | $\overline{\left(S_{\text {inv }} / V P\right) / N P} \overline{N P^{\uparrow}}$ | $V P / N P$ |
| $S_{\text {whq/o* }}\left(S_{\text {inv }} / N P\right)$ |  | $S_{\text {inv }} / V P$ |  |
|  |  | $S_{\text {inv }} / N P$ |  |
|  |  | $S_{w h q}$ |  |

(8)

$$
\frac{\text { Which }}{\frac{\left(\frac{S_{w h q} / \diamond \star}{}(S \backslash N P)\right) / N}{\frac{\text { woman }}{N}} \frac{\text { bought Ulysses? }}{S \backslash N P}} S_{\text {whq/o丸 }}(S \backslash N P) \text { Shq } \gg
$$

Crucially, the inclusion of raised (cased) argument types in the morpholexicon and rules of functional composition in syntax means that object Whquestions are immediately predicted to be unbounded-that is, an unbounded number of embeddings may separate the wh-element from the verb it depends on.
(9)


Once again, it emphasised that this mechanism for establishing unbounded dependency is quite different from the notationally similar gap-feature-passing mechanism of GPSG and HPSG. Whereas GPSG and HPSG pass an NP-gap feature from What to bought at each production in the derivation, the composition of $X / Y$ and $Y / Z$ to yield $X / Z$ is entirely independent of the presence elsewhere of a gap of type $Z$-or indeed of a gap of any type, as inspection of the many applications of composition in (9) and (??) will make clear.

In fact, the establishment of the dependency in the last step of (9) is more

[^1]akin to the uniform "swoop" account of movement of Ross (1967) than it is like the standard punctuated successive cyclic movement theory, much less the "hyper-cyclic" feature-passing at every rule-application of G/HPSG or Neeleman and van de Koot (2010), a point to which .

When more than one $W h$-element is questioned, then only one of them is fronted. (Chomsky, 1973) noted the "superiority effect" in English illustrated by the following apparent asymmetry: ${ }^{6}$
(10) Who bought what?
(11) \#What did who buy?

The possibility of in situ wh-elements in examples like (10) is accounted for if they are allowed to have order-preserving in situ Wh-question-forming categories like the one illustrated for what in the following derivation:

In the absence of a further in situ category $\left(S_{\text {inv }} / V P\right) \backslash\left(\left(S_{\text {inv }} / V P\right) / N P\right)$ for nominative who, raised over the inverting auxiliary, (11) cannot be derived. However, the inclusion of such a category would immediately overgenerate in situ $W h$-questions like the following with the meaning as (5):
(13) *Did who buy Ulysses?

There could not, therefore, be a language that just like English except for including that further category for who. The superiority asymmetry seems to be a side-effect of other idiosyncrasies of the English question-formation system. In other languages allowing multiple Wh-elements, superiority effects are not forced, as Bošković (2002) has shown for Slavic languages.

In Japanese, all wh-question forming expressions are in situ. Thus we have:

[^2]\[

$$
\begin{aligned}
& \text { "What did John buy?" }
\end{aligned}
$$
\]

It should be remarked that the English fronting Wh-question-forming categories given above, like all fronting categories, are semantically type-raised, but are not order-preserving, unlike the above Japanese $w h$-item. They conform to the Combinatory Projection Principle (CPP, (5) of chapter (2)) by marking their result $S_{w h q}$ as a different type of clause from the result $S$ of the the functor that they are raised over.

### 9.2 Indirect questions

Embedded or "indirect" questions of the kind already seen as (15) occur as the complement to verbs like "ask", "know", and "wonder":


Interestingly, indirect multiple wh-questions do not appear to suffer from superiority effects (or at least they are much reduced), supporting the earlier suggestion that such effects arise from other idiosyncrasies of English question forms, in particular do-support:
(16) a. I asked which woman bought which book.
b. I asked which book which woman bought.

### 9.3 Topicalization

Topicalization, as in (17), requires a "fronting" category similar to that of a $w h$-question or indirect question item such as what, as in (18):
(17) This movie, I like.
(18) a. What $:=S_{w h q / o \star}\left(S_{\text {inv }} / N P\right): \lambda p \lambda x \cdot[Q] p x$
b. what $:=S_{w h} /_{o \star}(S / N P): \lambda p \lambda n \lambda x$. $[Q] p x$
c. This movie $:=S_{\text {top }} l_{\circ \star}(S / N P): \lambda p \cdot p($ thismovie $) \wedge$ topic (thismovie)

Since more or less any argument can be topicalized, this is a further source of lexical ambiguity in arguments and their specifiers such as determiners and prepositions. As in the case of cased raised types, we can think of this either in terms of ambiguity "compiled out" into the lexicon, or in terms of underspecified categories like determiners $N P^{\uparrow} / N$. Since topicalization is a root transformation and can only occur in sentence intitial position in English, disambiguation by the processor is easy in either case.

Such raised categories are again order-changing, and again mark their result as a clause-type distinct from $S$, in conformity to CPP, (5) of chapter 2. Languages are free to include them or not, as we saw for the English "superiority effect" contrast in (10) and (11). Similarly, French lacks any equivalent of the topicalization category (18c), although in other respects its wh-constructions are quite similar to English.
The category (18c), lexicalized via determiners like this and certain intonational markers, maps similar residues onto English topicalized sentences like This movie, (I think that) you will like, with a side-effect asserting topic (thismovie), making this movie the discourse topic. ${ }^{7}$

If we can exploit the degrees of freedom in CCG to assign a distinct topicalizing category to phrases that act otherwise as in situ objects, as in (18c), then the same degree of freedom allows us to specific similar lexically distinct categories that are only allowed in fronted position. The following exclamatory construction (Huddleston and Pullum, 2002:918; Collins, 2005), which involve fronted elements like what a difference and How insensitive seems to involve such categories:
(19) a. What a difference a day makes.
b. How insensitive I must have seemed.

Thus, they are unbounded:
(20) a. What a difference she said that a day makes.
b. How insensitive I fear that I must have seemed.
7. Halliday, 1967b and Steedman, 2014 refer to topic in this sense as "theme", the latter offering a dynamic logic account of the side-effect.

However, unlike the apparently similar such a difference and so insensitive, they cannot occur as in situ objects:
(21) a. *A day makes what a difference.
b. *I must have seemed how insensitive.

Thus, they seem to involve the following additional categories for what and how:
(22) a. What $:=\left(S_{x c l} /(S / N P)\right) / N P: \lambda p . \operatorname{extremely}(p x)$
b. How $:=\left(S_{x c l} /(S / A P)\right) / A P: \lambda$ p.extremely $(p x)$

To the extent that both Topicalization and $W h$-questions resist embedding as complements in English, the fact can be captured by ensuring that no verb's subcategorization is compatible with the top and whq features they bear. However, this detail is language-specific, and contested even for English. ${ }^{8}$

### 9.4 Relativization

For English we can assume on semantic grounds that object relative pronouns have the following category, in which "|" is a slash whose value is either / or $\backslash:^{9}$
(23) who(m), which, that $:=\left(N_{a g r} b_{\star} N_{a g r}\right) / /_{\diamond \star}\left(S \mid N P_{a g r}\right): \lambda q \lambda n \lambda y . n y \wedge q y$

Like other " $w h$-moved" types, such relativized elements bear lexicalized non-order-preserving type-raised categories, which conform to the Combinatory projection principle (CPP), (5) of chapter 2, by yielding a type $N \backslash N$ that is distinct from $S$.
They support derivations like the following:

[^3]\[

$$
\begin{aligned}
& \text { (24) } \frac{\mathrm{A}}{N P_{\text {agr }}^{\text {b* }} N_{a g r}} \frac{\text { man }}{N_{3 s}} \frac{\text { that }}{\left(N_{\text {agr }} \partial_{\star} N_{a g r}\right) b_{b \star}\left(S / N P_{a g r}\right)} \frac{\text { Harry }}{N P_{3 s}^{\uparrow}} \frac{\text { detests }}{\left(S \backslash N P_{3 s}\right) / N P} \\
& \text { man }: \lambda q \lambda n \lambda y . n y \wedge q y: \frac{\lambda p . p \text { harry }: \lambda x \lambda y . \text { detests's }^{\prime} x y}{S / N P}>\mathbf{B} \\
& \xrightarrow[N_{\text {agr } \partial_{\star}} N_{\text {agr }}: \lambda n \lambda y . n y \wedge \text { detests } ' \text { y harry }]{\lambda \text { detests' } x \text { harry }} \\
& \xrightarrow[N P_{3 s}^{\uparrow}]{N_{3 s}: \lambda y \cdot \text { man }^{\prime} y \wedge \text { detests }^{\prime} y} \text { harry }
\end{aligned}
$$
\]

As in the case of $W h$-questions (9), the inclusion of type-raising (case) and composition in the grammar immediately predicts that the construction is unbounded:

$$
\begin{aligned}
& \text { (25) }
\end{aligned}
$$

$$
\begin{aligned}
& \xrightarrow{P_{3 s}^{\uparrow} N_{3 s} N_{\text {agr } \partial_{\star}} N_{\text {agr }}}<
\end{aligned}
$$

It is worth noting that the syntactic type of the relative pronoun resembles that of a type-raised NP determiner or generalized quantifier determiner, except that it is non-order-preserving and has a noun modifier $N \backslash N$ as its result, rather than $S$.
It should be apparent from derivations like (24) and (25) that it is only arguments like the object of detests or the complement of said that can be extracted or extracted out of. It follows that backward adjuncts (including many typeraised NPs) are predicted to be "islands" in the sense of Ross (1967).

The precise form of the present version of the slash-typing restrictions on rules (6) (which is slightly different from previous versions) is dictated by the interaction of the "Heavy Shift" (mediated by crossed composition) with extraction (mediated by harmonic composition) in examples like the following: ${ }^{10}$

[^4]

The inclusion of second-level composition (8) of chapter 4.1 allows related heavy-shifting derivations like the following:


The fact that in many dialects of English (including the author's), "inner" arguments of ditransitives can extract, means that we need further categories for the relative pronoun, schematized over ditransitive types, as in the derivation in floated figure 9.1a, crucially involving second-level composition $\mathbf{B}^{2}$, rule (8a) of chapter 4.1.

It is interesting to compare relativization in "structurally" cased English with Latin, because of the relatively unambiguous morphological case system of the latter. In floated figure 9.1b,unambiguous cased type-raised categories are spelled out, while $N P_{m}^{\uparrow}$ schematizes as usual over all cased masculine NP types.

English relative pronouns like (23) and Latin relative pronoun categories like quem, $\left(N P_{m, 3 s}^{\uparrow} \backslash N P_{m, 3 s}^{\uparrow}\right) /\left(S \backslash N P_{a c c}\right)$, above are examples of a general relative pronoun category of the form (28), mapping functions like $S \backslash N P$ into N and NP modifiers, in both cases agreeing in case with the residue of extraction: ${ }^{11}$
(28) $\left(N_{\text {agrl }} \mid N_{\text {agrl }}\right) /\left(S \mid N P_{a g r 2}\right): \lambda p \lambda n \lambda x . p x \wedge n x$

[^5]Figure 9.1:

The crossing dependencies between verbs and arguments in figure 9.1b arise from crossed composition reordering the nominative and accusative. Of course, this same reordering can be observed in Latin in non-relativized sentences, as in the scrambled version (7) of Chapter (??) of Figure 3.1 of Chapter 3.

Any such generalization will be subject to certain constraints on free order and combinatory rules that amount to saying that you can't do anything with an underspecified slash that you couldn't have done by proliferating distinct lexical entries with fully specified slashes (Baldridge 2002). In particular, the Combinatory Projection Principle (5) of chapter 2 requires that the interpretation of such underspecified slashes in the result of applying the composition rules (6) to such underspecified categories be confined to the alternatives allowed by those rules for the fully specified categories

As will be apparent from the Latin examples above, it is the inclusion of crossing composition rules in CCG that allows a limited amount of "scrambling" of word order, and hence allows non-peripheral arguments to extract. Together with the generalization of composition to boundedly polyvalent dependent functions such as $(Y / Z) / W$, the related valency limit on the coordinating categories, and with the limitations on the variety of such rules imposed by the Combinatory Projection Principle (5), it is the source of the "near-contextfree" expressive power of CCG (see Joshi et al. 1991; Koller and Kuhlmann 2009; Kuhlmann, Koller, and Satta 2010, 2015, and appendix C).

### 9.5 Bare relatives

We will assume that English bare relatives like $(30,31)$ are headed by a special relative pronoun like category for the subject NP and its determiners, by including the following categories, as well as the standard order-prederving ones, making them the head of a noun-modifier, $N \bigsqcup_{b+} N$ rather than $S$, and applying to a constituent of the type of a transitive verb rather than :

$$
\begin{aligned}
& \text { (29) I }::=\left(N b_{\star} N\right) /((S / N P) / N P): \lambda p \lambda n \lambda x . n x \wedge p x y \\
& \text { the }:=\left(\left(N b_{\star} N\right) /((S / N P) / N P)\right) / N: \lambda m \lambda p \lambda n \lambda x . n x \wedge \text { said }(p x)(\text { them })
\end{aligned}
$$

$$
\begin{align*}
\frac{\text { The }}{N P^{\uparrow} b_{\star} N} & \frac{\operatorname{man}}{N} \frac{\text { Harry }}{\frac{(N \backslash N) /((S \backslash N P) / N P)}{N \partial_{\star} N}} \frac{N}{(S \backslash N P) / N P} \tag{30}
\end{align*}>
$$



Since objects do not similarly head bare relatives in English, this analysis immediately captures the fact that bare subject relatives do not in general exist in English: ${ }^{12}$
(32) *This is the man likes Harry.

It may seem to proliferate nominal lexical category types recklessly. However, the ambiguity of the subject category is uniform and can be specified as a lexical rule, or be left unspecified as $N P^{\uparrow}$ in favor of (??) on the partial criterion of left-adjacent $N$ under the guidance of the parsing model. (This is in fact how such ambiguities are handled in practical wide coverage CCG parsers such as Hockenmaier and Steedman (2002a), Clark and Curran (2007b), and Lewis and Steedman (2014a) and the treebanks such as CCGbank that they are modeled on (Hockenmaier and Steedman, 2007).) ${ }^{13}$

A related analysis offers itself for VSO languages like Modern Welsh (Borsley et al., 2007) and Mayan (Ponvert, 2008), and for SOV languages like Japanese (Kuno, 1974, 1973a), in which there are no relative pronouns, and in that sense, all relative clauses are bare. The difference is that bare subject relatives are allowed in these languages, so that analysing them analogously to English bare relatives requires allowing all NP arguments of verbs to similarly head relative clauses.

Relative modifiers in these languages obey a generalization stated by Lehmann (1973):48, following Greenberg (1963), as follows:
(33) In VO languages, nominal modifiers such as relative, adjectival, and genitival expressions follow nouns; in OV languages they precede nouns.

For example, in Japanese we have:

[^6]a. [watakusi-ga kaita] hon
[ I wrote] book
$N / N \quad N$
"book (that) I wrote"
b. [hon-o kaita] hito
[book wrote] person
$N / N \quad N$
"person *(that) wrote a book"
(However, the object relativization in (34a) is only felicitous if the book is topic, either by virtue of $w a$-marking or by virtue of coreference with the topic, as in the following (Kuno, 1973b:234):)
(35) Kore-wa watakusi-ga kaita hon desu

This-TOP I-NOM wrote book is
"This is a book that I wrote"
See Kuroda (1965), Hasegawa (1985:n1), and below for further discussion, including the possiblity that Japanese relatives are mediated by pro-drop anaphora, as argued by Kuno (1973b). Accordingly, we remain for the moment uncommitted on exactly how $N \backslash N$ is derived in (34).)
In the case of Welsh, we can assume that for every raised NP type $N P^{\uparrow}$, schematized in (11) of chapter 3 as $\mathrm{T} \backslash(\mathrm{T} / N P)$, there is one or more corresponding types in which the argument $\mathrm{T} / N P$ has one fewer NP arguments and the result is (a function into) $N \backslash N$. These functions can only apply to the softmutated form of the verb, which we assume is distinguished by delivering $S_{\text {rel }}$ rather than $S$ as a result, since the soft-mutated form can head interrogatives in its own right. For example,
(36) a. dynes welodd cath

$$
\begin{aligned}
& \text { "woman (that) a cat saw" } \\
& \text { b. } \frac{\text { dynes }}{N: \text { welodd }} \frac{\text { gath }}{\frac{\left(S_{\text {rel }} / N P\right) / N P: \lambda y \lambda x . p a s t(\text { seexy })}{}\left(\frac{N \backslash N) \backslash\left(\left(S_{\text {rel }} / N P\right) / N P\right): \lambda p \lambda n \lambda x . n y \wedge p(\text { acat }) x}{N \backslash N: \lambda n \lambda x . n x \wedge \text { past }(\text { see }(\text { acat }) x)}<\right.} \\
& N: \lambda x . w o m a n x \wedge \text { past }(\text { see (acat) } x)
\end{aligned}
$$

"woman *(that) saw a cat"

Helpfully, as well as marking the relativized verb by soft mutation where applicable, as we saw in chapter4.1 Welsh distinguishes subjects from non-subjects, including those that head bare-relative clauses, by marking the latter with the soft-mutated form, a distinction that would be unmarked in English if it allowed bare subject-relativization.

It would be possible in principle to propose a mirror-image account of bare relativization for Japanese. However, we defer discussion of this question until we have dealt with some more complex relativization phenomena in English.

### 9.6 Embedded Subject Extraction

The fact that embedded subject extraction from that-complements is impossible in English, unlike extraction of other arguments of the verb, follows immediately from the present account, without further stipulations such as the Empty Category Principle (ECP) of Chomsky (1981) or the Generalized LeftBranch Condition (GLBC) or the related SLASH Termination Metarule of GPSG (Gazdar, 1981; Gazdar et al., 1985:161) and Trace Condition of HPSG (Pollard and Sag, 1994:172-3)
(37) a. a man who(m) $[I \text { think (that) }]_{S_{b *} S}[\text { Chester likes }]_{S / N P}$
b. *a man who(m) $[\mathrm{I} \text { think that }]_{S_{\delta_{*} S} S}[\text { likes Lester }]_{S \backslash N P}$
c. a man who(m) $[\mathrm{I} \text { think }]_{S_{b *} S}[\text { likes Lester }]_{S \backslash N P}$

While, as we have seen, extraction of an object (37a) is accomplished by the standard forward harmonic composition rule (6a), as in (25), extraction of a subject (37b) would require an unrestricted slash type $S^{\prime} / S$ in the category of verbs like "think" and/or the complementizer "that", allowing the forward crossed compostion rule (6c) to apply to yield *I think that likes Lester ${ }_{S \backslash N P}$.
However, as pointed out in $S S \& I$, if we made that change to the grammar of English to allow (37b), it would immediately also allow examples like the following:
(38) *I Chester think that likes Lester.

For the same reason, we cannot capture the exceptional cases of subject extraction that English allows for the small class of verbs that take bare complements, as in (37c), by making the appropriate entry for think compatible with both varieties of forward composition. Instead, we must add the following exceptional category for the bare-complement verbs:
(39) think $:=\left((S \backslash N P) / N P_{+W H, a g r}\right) /\left(S \backslash N P_{a g r}\right): \lambda p \lambda x \lambda y \cdot \operatorname{think}(p x) y$

This category is exceptional, in the sense that it violates the Projective Dependency Principle (1) (PDP) of chapter 2, in being a category that supports extracted but not in situ arguments. That is to say, it supports derivations like floated figure 9.2 for (37c), in which for the first time the agreement features on the English relative pronoun category(23) do important work.

The reason the extracting argument $/ N P_{+W H, a g r}$ has to be rightward, will become apparent in the next chapter 11 , in connection with example (7), where it supports across-the-board extraction of accusative arguments.

But it does not support a corresponding derivation with an "exceptionally case-marked" in situ accusative, because all lexical NPs are $-W H$ :
(40) a. *I think likes Lester my very heavy friend b. *I think him likes Lester.

The reader is directed to $S S \& I$ for further details, where it is shown, following Kayne 1983, that certain predictions that must follow if such PDPexceptional categories are allowed-for example, that similar constructions might exist that only allow extracted subjects, and disallow those in situ-are confirmed in French and English. For example, we have the following asymmetry in English:
(41) a. A man that I assure you to be a genius
b. *I assure you him to be a genius

The earlier work points out that the existence of dialects (including this author's) in which the following asymmetry is exhibited for the non-nominative relative pronoun whom is support for the analysis of embedded subject extraction as essentially accusative, an account which bears some relation to the account of Chomsky (1981) in terms of "antecedent government" of subject traces.
(42) a. A man who likes me
b. *A man whom likes me
c. A man who/whom I think likes me

See É. Kiss (1991) for discussion of related asymmetries in Hungarian.
If we can assume a special complement-subject-extracting category for certain verbs then we must predict that the same degree of freedom could be exploited for other categories taking bare S complements, in particular complementizers themselves, in other languages if not in English.

French provides a case in point. First, it strictly distinguishes the nominative

| man | who | I | think | likes | Lester |
| :---: | :---: | :---: | :---: | :---: | :---: |
| $\begin{aligned} & N_{3 s} \\ & : m a n \end{aligned}$ | $\begin{gathered} \left(N_{a g r} \backslash N_{a g r}\right) /\left(S / N P_{+W H, a g r}\right) \\ \lambda p \lambda n \lambda x . p x \wedge n x \end{gathered}$ | $\begin{gathered} \overline{N P_{l s}^{\uparrow}} \\ \lambda p . p m e \end{gathered}$ | $\begin{gathered} \left((S \backslash N P) / N P_{+} \text {WH,agr }\right) /\left(S \backslash N P_{\text {agr }}\right) \\ \lambda p \lambda x \lambda y \cdot \operatorname{think}(p x) y \end{gathered}$ | $\begin{aligned} & \left(S \backslash N P_{3 s}\right) / N P \\ & \lambda x \lambda y . l i k e s x y \end{aligned}$ | $\begin{gathered} N P^{\uparrow} \\ \text { lester } \end{gathered}$ |
|  |  | $\overline{\left(S / N P_{+}\right.}$ | $\underset{(, a g r)}{ }) /\left(S \backslash N P_{a g r}\right): \lambda p \lambda x \cdot \operatorname{think}(p x) m e$ |  | keslestery |
|  |  |  | $S / N P_{+W H, 3 s}: \lambda x$.think (likesle | terx) me |  |
|  | $N_{3 s} \backslash N_{3 s}: \lambda n \lambda x$.think (likeslester $x$ ) $m e^{\prime} \wedge n x$ |  |  |  |  |
|  | $N_{3 s}: \lambda x$.th | (likesle | x) $m e^{\prime} \wedge \operatorname{man} x$ |  |  |

Figure 9.2:
relative pronoun qui from accusative que, as in the following contrast:
(43) a. l'homme *qui/que les oiseaux voient ("the man that the birds see')
b. l'homme qui/*que voit les oiseaux ("the man that sees the birds')

Second, French has a lexically distinct complementiser qui that allows only embedded subject extraction for all verbs taking complements of that type:
(44) l'homme que je t'ai dit qui/*que voit les oiseaux" ("*the man that I told you that sees the birds").

Thus where Englsh has the categories in (45), French has the categories (46): ${ }^{14}$
(45) a. who, that $:=\left(N_{\text {agr }} \backslash N_{\text {agr }}\right) /\left(S \backslash N P_{a g r}\right)$
b. whom, who, that $:=\left(N_{\text {agr }} \backslash N_{a g r}\right) /\left(S / N P_{\text {agr }}\right)$
c. that $:=S^{\prime} / S$
d. think $:=\left((S \backslash N P) / N P_{+W H, a g r}\right) /\left(S \backslash N P_{a g r}\right)$
(46) a. qui := $\left(N_{a g r} \backslash N_{a g r}\right) /\left(S \backslash N P_{a g r}\right)$
b. que $:=\left(N_{\text {agr }} \backslash N_{\text {agr }}\right) /\left(S / N P_{\text {agr }}\right)$
c. que $:=S^{\prime} / S$
d. qui $:=\left(S^{\prime} / N P_{+W H, a g r}\right) /\left(S \backslash N P_{a g r}\right)$

Thus, French displays many of the alternative choices implicit within the degrees of freedom exploited in accounting for the idiosyncrasies of English in the possibiites allowed for subject extractions from tensed complements. The prediction is that the many other fixed word-order verb-medial languages that exhibit similar general constraints on subject extraction will be found the allow similar exceptions within the same degrees of freedom.

The fact that French does so in a way that allows embedded subjects to extract for all verbs taking qui-complements raises the possibility that there might be dialects or idiolects of English including a subject extracting category like (46d) for the complementizer that, as is suggested by the work of Sobin (1987) and Cowart (1997)—see Pesetsky (2017) for discussion.

The related question of how the child language learner can learn such lexicalized exceptions, rather than engendering collapse of the fixed word-order property by induction of over-general category types, as we noted would be appropriate in a veriable word-order language like Hebrew, which does allow

[^7]general subject extraction, is deferred until appendix B.
The *that-t phenomenon has been claimed above to be a consequence of English rigid word order and the differential directionality of the subject argument of the verb in SVO languages. It is therefore unsurprising to find that parallel constraints do not apply to verb-initial and verb-final languages and constructions. Thus in Italian, the subjects of exactly those verbs that allow subject inversion also allow embedded subject extraction (Perlmutter, 1971; Rizzi, 1982, 1990b; Ishii, 2004):
(47) a. Gianni abbia telephonato. ("Gianni has telephoned")
b. Abbia telephonato Gianni. ("It is Gianni who has telephoned")
c. Chi credi che abbia telephonato? Who.nOM think.2SG that has.PRES telephoned.PPL
("*Who do you think that has telephoned?")
This phenomenon is accounted for on the assumption that verbs like abbia have the VXS category $(S / N P) / V P_{p p l}$ as well as the standard SVO category, allowing subject extraction in (47c) via harmonic composition: ${ }^{15}$
\[

$$
\begin{gather*}
\frac{\text { Chi }}{S_{w h q} /(S / N P)}  \tag{48}\\
\frac{\begin{array}{c}
\text { credi che } \\
\frac{S / S^{\prime}}{S^{\prime} / S} \\
S / S \\
\hline \mathbf{B}
\end{array} \frac{\left(\overline{S / N P) / V P_{p p l}} \frac{\text { abbia }}{S / N P}\right.}{\frac{S / N P}{V P_{p p l}}}>\text { telephonato? }}{S_{w h q}}>\text { B }
\end{gather*}
$$
\]

Support for this explanation comes from the fact that VSO Welsh also allows embedded subjects to extract, assuming the account of Welsh relative clauses as headed by root NP categories, parallel to English bare relatives, introduced in section 9.5.
15. Modern Hebrew, which is generally regarded as SVO but both allows post-vverbal subjects and subject-extraction from complements, seems to be a similar case to Italian (Amir Zeldes p.c.).

| dynes/ddynes | (*a) wyddai | Gwyn | (y) gweliff | gath |
| :---: | :---: | :---: | :---: | :---: |
| $\begin{gathered} \hline N_{+f s} \\ \text { : woman } \end{gathered}$ | $\left(S_{\text {rel }} / S\right) / N P$ | $\left(\overline{\left.(N \backslash N) /\left(S_{\text {rel }} / N P\right)\right) \backslash\left(\left(S_{\text {rel }} / S\right) / N P\right)}\right.$ | (S/NP)/NP | $\overline{\left(S_{\text {rel }} / N P\right) \backslash((S / N P) / N P)}$ |
|  | : $\lambda y \lambda$ s.past (knowsy) | $: \lambda p \lambda q \lambda n \lambda x . n x \wedge p(q x) g w y n$ | : $\lambda y \lambda \lambda x$ will (seexy) | : $\lambda p \lambda \lambda y . p(a c a t) y$ |
|  |  | $\langle N) /\left(S_{\text {rel }} / N P\right)$ |  | $S_{\text {rel }} / N P$ |
|  |  |  |  |  |
|  | $N \backslash N: \lambda n \lambda x . n x \wedge$ past (know(will (see (a cat)x)) gwyn) |  |  |  |
|  | $N_{+f s}: \lambda$ x.woman | past (know (will (see (a cat)x) gw |  |  |

("woman that Gwyn knew will see a cat")
(50)

("woman that Gwyn knew a cat will see")
Again, it is helpful that Welsh disambiguates the verb and the residual NP in embedded $S_{\text {rel }} / N P$ by soft-mutation. The subscript rel is a purely syntactic feature that limits overgeneralization. We shall see it again in the relative constructions of the other Celtic languages Scots Gaelic and Modern Irish. Note that the fragment "wyddai Gwyn" ("Gwyn knew") has the same category as the English object relative pronoun, while the remainder of the relative clause "(gwybodd e) gweliff cath", "(he knew) a cat will see", in which the verb(s) have the standard unmuted form, has the same category as the corresponding residues of English object relativization.

As predicted, the verb-final language Japanese also shows no assymetry in extraction of embedded subjects and objects. In particular it too allows embedded subject-relativization under the conditions noted in connection with (35) (Kuroda, 1965; Kuno, 1973b):
(51) Kore-wa John-ga hon-o kaita to itta hito desu This-sc top John-NOM book-ACC wrote that said person is
"This is the person who John said (*that) wrote the book"
However, we continue to defer discussion of Japanese relativeization.
There is considerable variation across Germanic dialects and individual
speakers as to whether they allow any extraction of embedded arguments, with Northern German speakers tending to disallow any extraction from embedded clauses, while Southern speakers tend to allow extraction of both subjects and objects from embedded clauses. This observation suggests that in the former dialects complementizers such as "daß" have the same category as Irish "go", while in the latter they have the same category as Irish "a(L)".

Bresnan (1977:173;194,n.7), Culicover (1993), and Browning (1996:237,n.1) discuss some cases of English embedded subject extraction where the inclusion of a sentential adjunct adjunct seems to facilitate subject extractions that are otherwise disallowed. The status of many of these examples is rather unclear, as Bresnan and Browning point out, but the most convincing examples involve the negative adverbials that precipitate inversion, as in (52b).
(52) a. a person who(m) I said that under no circumstances would run for any public office
b. I said that under no circumstances would this person run for any public office.

Such examples are parallel to the possibility of complement subject extraction in Welsh discussed above, since under present assumptions negative fronted items require an inverting VSX verb such as the following, which do not require crossed composition for the subject to be accessible:
(53) would $:=\left(S_{\text {inv }} / V P\right) / N P$

Other extraction-facilitating adjuncts discussed by these authors such as (54b) are harder to explain:
(54) a. a person who(m) I said that *(in my opinion) was unfit for any public office.
b. a person who(m) I wonder whether *(in your opinion) is fit for public office.

SS\&I:61 suggests that such adverbials carry a wh-extraction only category $\left(S / N P_{+W H, a g r}\right) /\left(S \backslash N P_{\text {agr }}\right.$, a suggestion that is supported by the fact that the effect seems to hold for verbs that do not support bare complements, as in (54b).

Many analyses of the constraints on complement subject-extraction and the exceptions to those constraints have been proposed, and are reviewed by Pesetsky (2017). It seems fair to say that none of them are entirely satisfactory. Under the present account, the possibility of asymmetry in extractability of subjects and objects in rigid-order SVO languages like English follows from their lexical specification in their governing category as different in direction-
ality, from which it follows that distinct combinatory rules must apply. Object extraction requires rightward harmonic composition of the complement-taking $V P / S$ into the tensed verb category, so it is potentially allowed. Embedded subject extraction would require forward crossed composition into the same category. Since that is a different rule a language is free to independently disallow that extraction. What is more, allowing such crossed composition would immediately allow very free word-order, so we would no longer be able to talk of the language as SVO in the first place.

Conversely, the tendency towards symmetry of embedded subject and object extraction in verb initial (VSO) and verb final (SOV) languages follows from their lexical specification in their governing category as the same in directionality, from which it follows that the same combinatory rules must apply to them. Thus if a rule of composition allows extraction of objects, then it must apply to subjects. It follow that such languages can exhibit symmetry in extraction, either allowing or disallowing both. In fact, the only way that such languages could exhibit extraction asymmetry would be via the agreement system, say by restricting relative pronouns to nominative, absolutive, etc. agreement, as in Latin (Figure 9.1, Kennedy, 1882:§330), and other cases discussed by Keenan and Comrie (1977).

### 9.7 Island Constraints

The various types of long-range dependency considered in the preceding sections have one characteristic in common: In every case the displaced element is an argument, such as $N P$, and the domain with respect to which it is displaced is also an argument, such as $V P$, or $S / N P$, itself composed from heads and complements of arguments. The displaced element may bear its standard, case-raised category, as in Japanese scrambling. Or it may be a special category-changing type, like an English topic. But in all cases it is a (raised) argument. ${ }^{16}$

If a category is an adjunct, such as the adverbials "naked" and "whistling Dixie", $V P \backslash V P$, then it is defined as selecting, rather than selected-for, thereby rendering inapplicable all the mechanisms above for extraction. Moreover, incomplete adjuncts such as "whistling" $(V P \backslash V P) / N P$ cannot combine to the

[^8]left with $V P$ in advance of rightward with $N P .{ }^{17}$
Under present assumptions adjuncts are therefore predicted to be islands to extraction, as observed by Huang, 1982a:505, following Cattell (1976): ${ }^{18}$
(55) a. *How [do you regret that Sally met Harry? $]_{S_{i n v}}$
b. *Who did Harry [file the reports] $]_{V P}[\text { without telling }]_{(V P \backslash V P) / N P}$ ?
c. *Who do you know a $[\text { man }]_{N}[\text { that met }]_{(N \backslash N) / N P}$ ?

Since we have type-raised all arguments including indirect questions, in most cases excluding the corresponding unraised category from the lexicon, and type-raised categories are VP-adjunct-like, in the sense that they select verbs and verb-phrases, as in $V P \mid(V P \mid X)$ etc, we correctly predict that raised arguments including $N P^{\uparrow}$ and indirect questions are also islands:
> a. *What did you [doubt] $]_{V P / N P}$ the claim ${ }_{N}$ [that he has read $]_{(N \backslash N) / N P}$ ?
> b. *Which woods do you think you $[\text { know }]_{V P / S_{i q}}$ [who $\mathrm{owns}_{\left(V P \backslash\left(V P / S_{\text {iq }}\right)\right) / N P}$ ?

The possibility of explaining island effects in this way is another desirable consequence of specifying linear correspondence in the lexicon.
Certain constructions that look like extractions from NP and PP like the following possibly arise from lexicalized multi-word expressions or verb-particle constructions, predicting that such expressions as subjects are islands like all arguments, because noun-modifiers are adjuncts $N \backslash N:{ }^{19}$
(57) a. Who did they take advantage of?
b. What will you paint a picture of?

In English, S, $\bar{S}$, 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 like the following, analogous to (??), suggests that they also bear adjunct or raised categories: ${ }^{20}$

[^9](58) I will tell ${ }_{(V P / S) / N P}$ [[[Donald] [(that) he is fired]] and [[Ronald] [(that) he is hired] $]]_{V P \backslash((V P / S) / N P)}$.

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 typeraised categories, making them islands, as appears to be the case for daßcomplements in many dialects of German, or including different complement types, some of which are adjuncts, and some subcategorized-for arguments, as is the case for wh/that-complements in English:
(59) 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 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, 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:
(60) a. *What tune does John work whistling?
b. What tune did John drive Mary crazy whistling?

Example (60a) shows that VP-modifiers like "whistling Dixie" are not subcategorized-for by predicates like "work" $V P$ : the only way such modifiers can combine with them is as adjuncts $V P \backslash V P$, which are islands. However, (60b), shows that they also carry the argument category $V P_{\text {ing }}$, allowing extraction past subcategorizing verbs, as in "What tune is John whistling?". In the case of (60b), this implies a category for "drive" and related causatives like "make" of $\left(\left(V P / V P_{\text {ing }}\right) / X P_{\text {pred }_{i d v}}\right) / N P$, subcategorizing for $V P_{\text {ing }}$ and allowing (60b) by composition.

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 with suitable matrix verbs. 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", under which John merely happens to whistle that tune while he does whatever it is that actually drives her crazy.

The exact conditions under which weak island ambiguities are resolved in favor of the complement to permit extractions depend upon the matrix-verb's subcategorization, the parsing model, and/or world knowledge, essentially as proposed in neural-computational terms by Dowty (2003) and in eventsemantic terms by Truswell, rather than upon syntax alone.

Kuno (1973b) noted that Japanese allowed certain relative clauses that appear to violate such island constraints. An example is the following: ${ }^{21}$
(61) [[katteita] inu-ga sinde simatta] kodomo
[[kept] dog-NOM die-PERF] child
"child ${ }_{i}$ who the dog that (\#he ${ }_{i}$ ) kept died."
This possibility is not predicted by the analysis of Japanese relatives floated in section 9.5 , in which relative clauses were hypothesised to be bare adjuncts $N / N$, and has been analysed extensively by Hasegawa (1985) and Richards (2001) in terms of very powerful movement theories that cannot be simulated in CCG terms. ${ }^{22}$

However, Kuno and Hasegawa point out that such examples are only allowed under the condition that the apparently island-violating extraction (here, headed by the child) is the subject, theme or topic of the inner clause, ("kept"), as it is here. Since Japanese is both topic-prominent and pro-drop, it seems possible that the subject long-range dependency in question is anaphorically mediated, as it is by a resumptive pronoun in the corresponding English given above. Indeed it seems possible that all Japanese relativization is mediated by pro-drop, hence an anaphoric rather than syntactic phenomenon. (Kuno's own "topic deletion" account is compatible with this proposal.)

### 9.8 Preposition Stranding and Relativization

The possibility exhibited by English and some Celtic and Scandinavian languages of extracting complements of prepositions, as in the following, is crosslinguistically extremely rare (van Riemsdijk, 1978):

[^10](62) a. Who did you buy the bicycle from?
b. I [sold a book to and bought a bicycle from] a very nice man,
c. I took [the bus to and the train from] 30th Street station.

No other Germanic or Latinate language, including Latin itself, strands adpositions with anything like this generality. Prepositions in such languages behave much like determiners, suggesting that in most languages they are case-like operators turning their complements into type-raised categories, which, being adjuncts, cannot be extracted out of. For example, the following seems to be the category of the Dutch preposition op, "on/in", as the specifier of a verbal argument:
(63) op $:=P P_{o p}^{\uparrow} / N P$

Like all the prepositions considered here, op has another category as the head of a sentential adverbial adjunct:
(64) op $:=(V P \backslash V P) / N P$
as well as that of a particle in constructions like opbellen "ring up".
In English, by contrast, argument prepositions like "to" arguably do not bear any category related to (63) in alternation with the adjunct-heading category (64). Instead, they seem to bear ab adjunct-particle-like type-changing category like the following, restricted in the lexicon by the $\diamond \star$ slash-type.
(65) to $:=(V P / o \star A P) b_{\star}\left(V P / P P_{t o}\right)$

This category allows the following derivation:

It also allows the following alternative derivation, among others, for the same sentence:

Coordinate sentences like the following are thereby allowed as a form of argument-adjunct cluster coordination, introduced in Chapter 4.1:
(68) I will give flowers [to Henry and to Sam] ${ }_{V P\rangle_{\Delta \star}\left(V P / P P_{t o}\right)}$.

The category (65) allows preposition-stranding extraction:

(70) la


As predicted by the earlier account of argument/adjunct cluster coordination in chapter 4.1, preposition-stranding sequences like "to tomorrow" can freely coordinate, as in the following example:
(71) A land that I will travel [to today and from tomorrow] $]_{(V P / N P)_{ぬ \star}}\left(V P / P P_{t o}\right)$

However, the $\diamond$ modality on (65) prevents violation of the constraint that in the heyday of Constraints on Transformations Kuno (1973a) jokily entitled the "Clause Non-final Incomplete Constituent Constraint (CNFICC)" on preposition stranding out of the Heavy NP Shift construction analysed in example (9) of Chapter ??:
*a man who I will

$$
\begin{equation*}
\frac{\frac{\text { give }}{\left(\frac { 1 } { V P / P P _ { t o } ) / N P } \left(\frac{\text { to }}{V P / N P) d_{\star}\left(V P / P P_{t o}\right)}\right.\right.} \text { that very heavy book }}{* * *}<\mathbf{B}_{\times} \tag{72}
\end{equation*}
$$

The modality that prevents preposition stranding of this category via crossed composition would also block Heavy NP Shift itself contrary to fact.. It follows that Heavy NP Shift over PP must involve the other, adverbial, category of PP and syntactically transitive sent, analogous to the derivation in (9), where the latter category semantically includes an argument that we write $s k_{\text {recipient }}$ which is anaphoric to something with the property of being a recipient, the property which the adjunct predicates of Harry: ${ }^{23}$
$S:$ recipient harry $\wedge$ sent $\left.s k_{\text {recipient }} s k_{\text {veryheavybook }} i\right)$
The above analysis of stranding prepositions resembles a lexicalized version of Hornstein and Weinberg's and Kayne's (1981) "reanalysis" accounts of preposition stranding. The paradigm discussed in this section was also a major movation for Pesetsky's 1995:176 notion of "cascade" structure-see Phillips $(1996,2003)$ for an extended comparison between Pesetsky's approach and CCG. We shall see in chapter 11 further evidence in support of this account of English preposition stranding from adjunct/argument cluster coordination. ${ }^{24}$

### 9.9 On "Remnant Movement"

Both Heavy NP Shift and crossing dependencies in serial verb constructions have been analysed under the movement theory in terms of remnant movement-that is, the movement of constituents that already include a trace as a result of some other movement. The effect of crossed composition is to accept exactly those word orders that if derived from German- style verb-final embeddings to general the "English" orders available in Dutch and Hungarian, give the appearance of moving clauses with gaps in them.
23. We defer discussion of such terms until chapter 13 . We could achieve the same effect with transitive semantics and a Davidsonian event variable, at the cost of some notational clutter. 24. Contrary to the prodictions of the above analysis, Pesetsky claims (1995:249) that Heavy Shift is clause bounded, claiming an asymmetry in acceptability with leftward extraction in examples like the following:
(i) a. What gift will Mary [claim that she gave to Harry] ${ }_{(S \backslash N P) / N P} \operatorname{TOMORROW}_{(S \backslash N P) \backslash(S \backslash N P)}$ ?
b. Mary [will claim that she gave to Harry] $]_{(S \backslash N P) / N P}$ tomorrow $_{(S \backslash N P) \backslash(S \backslash N P)}$ a very heavy BOOK.
However, any such asymmetry seems to arise from the greater demands that heavy NP shift makes on the preceding context: (ib) seems perfectly acceptable as an answer to the question (ia).


Figure 9.3:

The availability of crossed composition is also both necessary and sufficient to support the analysis in figure 9.3 of "verb fronting" or verb topicalization in German, which has been held to provide major evidence for rules of remnant movement or "discontinuous constituency" (Reape 1994; Müller 1998; De Kuthy and Meurers 2001; see Müller 2007; Wechsler 2015:234-6 for discussion and a literature review).

This analysis, in which the ditransitive infinitival main verb, like all topics, receives a lexicalized exotypic order-changing type-raised function into topicalized sentences $S_{\text {top }}$ (cf. English 18b), is equivalent to that of Nerbonne (1986) and Johnson (1986), which as they point out are essentially combinatory-categorial. ${ }^{25}$

The following asymmetry discussed by Johnson, which is left as an exercise, is also predicted:
(74) a. Erzählen können hat er seiner Tochter ein Märchen.
b. *Können hat er seiner Tochter ein Märchen erzählen.

### 9.10 Tough-movement

Consider the examples in the following alternation:
(75) a. To (imagine that we could) please John is easy
b. It is easy to (imagine that we could) please John.
c. John is easy to (imagine that we could) please.

The first two sentences $(a, b)$ are just an example of the copular propositional subject construction and its alternation with extraposition or dislocation that was discussed in section 7.4:
(76) a. Being green isn't easy/much fun/a cakewalk.
b. To err is human/embarassing/a pity
c. That they won is unfortunate/surprising/a triumph.
(77) a. It isn't easy/much fun/a cakewalk being green.
b. It is human/embarassing/a pity to err.
c. It is unfortunate/surprising/a triumph that they won.

[^11]The categories for easy are as follows: ${ }^{26}$
(78) a. easy $:=X P_{\text {pred }_{i d v}}: \lambda y$.easyy
b. easy $:=V P_{\text {pred }_{\text {idv }}} /\left(V P_{\text {to }} / N P\right): \lambda p \lambda x$.easy ( pxone)

As in the case of the short passives discussed in the last chapter, the constant one is a placeholder representing an arbitrary agent whose semantic nature need not concern us here, but to which we will return in section 13.6.3:

The derivation of $(75 \mathrm{c})$ is then as in (79):

This analysis in essence follows those of Carpenter (1992) and Jacobson (1992a), who point out that functional composition allows the infinitival transitive verb to be an unboundedly large fragment such as to imagine that we could please, accounting for the unbounded character of the dependency involved, while maintaining the Principle of Combinatory Projection. ${ }^{27}$

The limitation of the tough construction to nesting its dependency in the minimum pair shown in figure 9.4a,b (Chomsky, 1977b) follows immediately from the lexical analysis and the mechanism for extracting inner arguments (cf. figure 9.1a):

The intended reading with sonatas played upon the violin cannot be obtained from figure 9.3b, because the stranded preposition upon cannot combine with are easy to play until the relative pronoun has combined with the latter. However, that cannot happen until violins has combined, and the only way for that to happen is for the latter to have the category of a subject, forcing the unintended reading with \#the violin played upon sonatas. Nor is there any other assignment of legal CCG categories that will allow the intended meaning to be derived

Thus, CCG offers a solution to the problem of an asymmetry which appears to remain open or stipulative in solution in other theories of grammar including

[^12]

Figure 9.4:

G/HPSG and MP (Chomsky, 1977b; Gazdar et al., 1985:150-153; Pollard and Sag, 1994:169; Hornstein, 2001; Hicks, 2009),

### 9.11 Pied-Piping and Wh-In-Situ

Following Morrill 1994, 1995 and Steedman 1987, TS:89-91 and SS\&I:5051 propose an extra lexical category for pied-piping wh-items such as which and who(m) in NPs such as reports [the height of the lettering on the covers of which] the government prescribes, which can under present assumptions concerning the type-raising of NPs be written as in (80):
(80) who $(\mathrm{m})$, which $:=((N \backslash N) /(S \mid N P)) \backslash\left(N P^{\uparrow} / N P\right): \lambda p \lambda q \lambda r \lambda x . q(p x) \wedge r x$

Apart from the fact that relativizers are functors into $N \backslash N$, rather than $S$, this category is simply that of a type-raised NP raised over functions into typeraised NPs-that is, $N P^{\uparrow} \backslash((S /(S \backslash N P)) / N P)$.

We know that composition can form constituents like the government prescribes and the height of the lettering on the covers of, since we can extract out of them:
(81) Reports which the government prescribes the height of the lettering on the covers of

So the category in (80) allows the following:
(82) $\left[[\text { Reports }]_{N} \text { [the height of the lettering on the covers of }\right]_{N P^{\uparrow} / N P}$ $[\text { which }]_{((N \backslash N) /(S / N P)) \backslash\left(N P^{\uparrow} / N P\right)}$ [the government prescribes $\left.]_{S / N P}\right]_{N}$

The pied-piping wh category $N P^{\uparrow} \backslash\left(N P^{\uparrow} / N P\right)$ is also the type we need for the related category of in situ wh-items like which reports in "quiz show" questions like the following:
(83) The government prescribes [[the height of the lettering on the covers of $]_{N P^{\uparrow} / N P}$ [which reports $\left.]_{N P^{\uparrow} \backslash\left(N P^{\uparrow} / N P\right)}\right]_{N P^{\uparrow}}$ ?

The availablity of a special-case "antecedent-controlled" subject extracting category (39) for bare-complement verbs like know, $\left(V P / N P_{+W H, a g r}\right) /\left(S \backslash N P_{\text {agr }}\right)$, does not support in situ complement subject $w h$-items for those verbs. In situ wh-embedded subjects are in all cases just as unacceptable as subject extractions over complementizers like that:
(84) a. *Who believes which candidate will win the election?
b. *Who believes that which candidate will win the election?

The in situ wh-elements discussed above are well known for being immune in other respects to island constraints, as shown in the following comparison for English:
(85) a. \#Which famous movie did Mary marry the man who directed?
b. Mary married $\left[\right.$ the ${ }_{N P^{\uparrow} / N}\left[\operatorname{man}_{N}\left[[\text { who directed }]_{(N \backslash N) / N P}\right.\right.$ [which famous movie $\left.\left.\left.]_{(N \backslash N) \backslash((N \backslash N) / N P)}\right]_{N \backslash N}\right] N\right]_{N P^{\uparrow}}$ ?

The in situ wh-item can have a category type-raised over the type $(N \backslash N) / N P$ of who directed, and apply to it to yield the noun-modifier who directed which musical $N \backslash N$, which then combines in the standard way with the noun man and determiner the to yield a raised object NP, which can combine with the matrix in the standard way to yield $S$. (The question illocutionary force appears to stem from intonational accent on the wh element.)
We noted earlier in section 9.5 that the same immunity from islands is known to apply quite generally in languages like Japanese, in which all overt whelements are normally in situ (Kuno, 1973b). For example:
(86) Mary-wa [[John-ni nani-o ageta $]_{S^{\prime}}$ hito-ni $]_{N P}$ atta-no Mary-TOP [[John-DAT what-ACC gave $]_{S^{\prime}}$ man-DAT $]_{N P}$ met-Q
'\#What did Mary meet the man who gave to John/Mary met the man who gave what to John??'

Choe (1987), Nishiguchi (1990), Pesetsky (1987), Watanabe (1992, 2001), Brody (1995), von Stechow (1996), and Richards (2000) have variously argued, against Huang (1982b) and Lasnik and Saito (1984), that in situ whconstructions escape island constraints because of the possibility of covert "large-scale" pied-piping of entire phrases like John-ni nani-o ageta hito-ni, the man who gave what to John.

Although the present account of pied-piping does not involve syntactic movement, those arguments, and in particular the "across-the-board" condition on the scopes of multiple in situ wh-items inside islands discussed by Watanabe and Richards, appear to support it.

In particular, as in the case of the basic English relative pronoun (23), the pied-piping category (80), lexicalizes exactly the same insight as covert largescale pied-piping movement, albeit statically at the level of lexical logical form, using only the standard variable-binding apparatus of the $\lambda$-calculus.

In the light of the homomorphic relation of prosody and intonation structure to syntax propsed in chapter 6 , it should be clear at this point that the theory of relativization proposed in this chapter is closely related to the insight
of Richards $(2010,2016)$ relating the scope of $w h$ to a contiguous prosodic phrase. The only difference is that under the present theory this identity holds for both wh in situ and wh-movement, with combinatory derivation providing the domain for both, as well as for the formation of intonational phrases discussed there.
That is simply to say that the domain of in-situ wh and pied-piping wh are both, like ordinary pronominal relativization, defined by combinatory constituency. The only difference between the two forms of $w h$-construction and the bare forms of relativization considered in section 9.5 above is the languagespecific lexical specification of wh-elements as either leftward- or rightwardcombining.

### 9.12 Celtic relativization

Interestingly, in terms of the distinction within transformationalist theory from the "swoop" theory of unbounded movement over multiple tensed boundaries as a single operation of Aspects to the "cyclic" theory of unbounded dependencY as multiple local movements within successive tensed domains introduced by Chomsky (1973), the present non-movement theory is closest to a swoop theory, since the wh-element combines with the entire combinatory residue in a single merger.
The resemblance of CCG to a swoop, rather than a cyclic, account of movement might appear to conflict with the claims of McCloskey (1979); Adger (2003) and Boeckx (2008) to have proven on the basis of a number of constructions in a number of typologically distinct languages that movement is necessarily cyclic, and not swooping.
Adger 2003:376-386 provides a convenient summary of the argument based on wh-constructions in Scottish Gaelic, which is essentially parallel to McCloskey, 1979:150 for Irish, and to Welsh, seen earlier.

All wh-questions in Gaelic include an overt complementizer "a":
(87) Cò an duine a tha thu a' pòsadh Who the man that are you ing- marry
"Which man are you marrying?"
When the wh-dependency is embedded, the " a " complementizer is obligatorily repeated:
(88) Cò an duine a bha thu ag ràdh *(a) bhuail i

Who the man that were you ing- say that hit she
"Which man were you saying that she hit?"
The " a " complementizer involved in long-range $w h$-constructions is in complementary distribution to the standard complementizer "gu(n)", which specifies the "dependent" form of its complement verb, here "do bhuail" ("hit"). The "gu(n)" complementizer acts as a barrier to extraction:
(89) *Cò an duine a bha thu ag ràdh gun do bhuail i Who the man that were you ing- say that prt hit she
"Which man were you saying that she hit?"
In general, Adger notes the following patterns for embedded complements:
(90) a. ...gu(n) $\ldots g u(n) \ldots$
(complementation)
b. *...gu(n) $\ldots a \ldots$
c. $\ldots a \ldots a \ldots t$
(wh-relativization)
d. $* \ldots a \ldots g u(n) \ldots t$

These authors claim that if the movement from "bhuail i" ("hit she") to the root "Có" were to take place as a unitary swoop, then we would have no way to prevent (89), since the movement would not be blocked by being unable to land at the illegal intermediate complementizer. Only if movement is cyclic can the moving element notice that its upward passage is blocked.
However, this claim is clearly theory-internal, and only makes sense on the assumption that unbounded dependency is literally mediated by a rule of displacement. We can assume instead that the relativizing complementizer "a" is a standard complementizer marking its clause as of the special type $S_{\text {rel }}$ that was first encountered in the analysis of Welsh bare relative clauses in section 9.5, while "gu(n)" is like Northern German "daß" the specifier of a type-raised island complement, implying the following lexical categories:
(91) a $\quad:=S_{\text {rel }} / S: \lambda s . s$
$\operatorname{gu}(\mathrm{n}):=\left(S \backslash\left(S / S^{\prime}\right)\right) / S_{\text {dep }}: \lambda s \lambda p . p s$
On the assumption that $W h$-questions in Gaelic depends upon the combination of a wh-element such as "Cò"of type $S_{w h q} /\left(S_{\text {rel }} / N P\right)$ with a constituent of type $S_{\text {rel }} / N P$ formed by combinatory composition, differing only from the corresponding English categories in specifying the relativized form in the residue, then the contrast between (88) and (89) can be captured without any explicitly
cyclic operations or any other kind of action-at-a-distance:

"Which man were you saying that she hit?"

"Which man were you saying that she hit?"
Relative clauses in Scots Gaelic also involve the relative complementizer "a" and exclude "gu(n)". However, as for other Celtic languages, 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).

We therefore make the same assumption as for Welsh, that one of the NP complements of the matrix verb is the head of the relative clause of type $N \backslash N$, with the additional assumption that the Scots "a" complementizer has a further category mapping that category to itself
(94) a $:=S_{\text {rel }} / S: \lambda s . s$

$$
:=(N \backslash N) /(N \backslash N): \lambda p \lambda n . p n
$$

We then have the following pattern of derivations for Gaelic long-range relativization (90c): ${ }^{28}$
28. It will be clear from the derivation why it is frequently possible to elide " $a$ ".

"man that he said that she hit"
Note that "thuirt e" ("said he") in the above carries the same category as a matrix English object relative pronoun. Embedded verbs like "bhuail" carry the standard independent $S$ type.)

By contrast, "gu(n)" acts as an island-inducing type-raised argument head, requiring a dependent $S_{d e p}$ complement: ${ }^{29}$
(96)

"He said that she hit him"
However, if the relative complementizer "a" intervenes no similar derivation of a complete complement is possible:

Similarly, embedded island barrier "gu(n)" blocks the formation of "said he that hit she", the argument of long range relativizer "a":

[^13]

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. ${ }^{30}$

### 9.13 Discussion

The combinatory treatment of the relative clause proposed here is syntactically mediated solely by combinatory reductions operating on adjacent typed constituents. CCG can therefore be seen as reducing the transformationalists' operation MOVE to (external) MERGE.

The long-range dependency at the heart of the construction is established via the lexical logical form $\lambda q \lambda n \lambda y \cdot q y \wedge n y$ of the relative pronoun (23), and in particular by the use there of the second-order variable $q$.
One might also view this second-order $\lambda$-term as formalizing some version of the "copy theory of movement," with the $\lambda$-bound variables doing the work of "copying" declaratively, at the level of static logical form. In that sense one could also see CCG as reducing "overt" syntactic movement to "covert" LF movement, as does the "Trace Conversion Rule" of Fox (2002).

However, there is no process of copying as such. This is simply the standard apparatus of binding a value to a variable that occurs, possible more than once, in a logical formula. The relation of the head noun and subordinate clause is defined once and for all, in the lexical logical form, which the combinatory syntax merely projects by adjacent merger onto sentential logical form. Any constraints on possible projections can only arise from the combinatorics of syntactic projection.

Exercise The pattern in Irish is similar to Gaelic, but complicated by the possibility of resumptive as well as wh-relativization. McCloskey (2002:193200) identifies the following more complex pattern, of which the first three are the most frequent: (See also McCloskey, 2017.)

[^14]
Try to extend the analysis of Scots Gaelic relatives in section 9.12 to Irish.


[^0]:    1. "The eternal silence of these infinite spaces terrifies me."
    2. Cf. $S P: 33$, where the same principle is called "Head Categorial Uniqueness".
[^1]:    5. We shall see in chapter 6 on spoken intonation that words with and without accent count as distinct lexical items, much as stems with and without tense morphology do.
[^2]:    6. The superiority effect is weak for some speakers, and the judgements correspondingly uncertain.
[^3]:    8. Miyagawa (2017) shows that embedded topicalization is possible under certain circumstances in Japanese.
    9. The significance of the agreement features on the category when it is / will become apparent when we consider subject extraction in section 9.6.
[^4]:    10. Details of agreement are suppressed to limit visual clutter.
[^5]:    11. See section 9.6 below for details of English embedded subject extraction.
[^6]:    12. They are occasionally found in poetic registers-see the epigraph to Chapter 7 for an example. 13. Another analysis is possible in which it is the verb that heads bare relative clauses, via categories like $(N \backslash N) / N P$ ) for the transitive. This also could be handled via underspecification and/or unary rules.
[^7]:    14. The agreement-passing variable feature-value $a g r$ is shown for once, since it matters for subject-extraction. Recall that lexical NPs bear the feature $-W H$, and cannot combine with $N P_{+W H}$.
[^8]:    16. The same holds for "remnant" movements, such as Germanic transitive verb-topicalization, as in "Essen wird er Æpfel" ("*Eat will he apples"), (Müller, 1996), since $V P \backslash N P$ is the argument of the German accusative NP, allowing the movement residue $S_{i n v} /(V P \backslash N P)$ to form, unlike the corresponding residue in English.
[^9]:    17. Other, that is, than via the parasitic gapping $\mathbf{S}$ substitution rules we have passed over in section ?? (Szabolcsi, 1983/1992; Steedman, 1987, 1996).
    18. Substituting "bridge" verbs like "think" for non-bridge "regret" considerably improves (55a), because bridge-verbs actually do optionally subcategorize for adjuncts.
    19. The relative weakness of "derived" subject islands such as unaccusatives found by Jurka (2013) and Polinsky, Gallo, Graff, Kravtchenko, Morgan, and Sturgeon (2013) is not explained in these terms. The latter authors, citing Chung and McCloskey (1983) and Kluender (1998), note definiteness effects that suggest that discourse processing factors may be at work here too.
    20. It seems likely that as adjuncts they are extraposed, with an anaphoric relation to an argument of verbs like "ask" and "tell" at the level of logical form.
[^10]:    21. We noted earlier that Japanese, being SOV, allows embedded subjects to extract, unlike English.
    22. Richards $(2000,2001,2002)$ relates such examples to the possibility of in situ multiple $w h$ elements, which are briefly discussed in non-movement terms in section 9.1 above.
[^11]:    25. The significance of non-standard constituents like seiner Tochter ein Mürchen $_{V P /\left(\left(V P \backslash N P_{\text {dat }}\right) \backslash N P_{\text {acc }}\right)}$ is discussed in section 11.4. Any corresponding verb topicalization in English *tell he will his daughter a story is prevented by the same independent features of English grammar that both contribute to its rigid word-order and in general prevent subject extraction, namely $\diamond$ modality on complement taking verbs including the modals (SP:53-4).
[^12]:    26. Further categories for easy are needed to capture the related sentences It is easy for us to . . . please John and John is easy for us to . . . please. See Jacobson 1992b for discussion.
    27. Jacobson points out that the analysis has earlier antecedents in unpublished work by Gazdar, and in Chomsky 1977a.
[^13]:    29. Semantics is suppressed as similar to English.
[^14]:    30. A related analysis to that of Gaelic above appears to be applicable to the Germanic "whcopying" phenomenon discussed by Felser (2004).
